JOURNAL OF DIFFERENTIAL EQUATIONS 32, 193–232 (1979)

A Finite Difference Technique for Solving Optimization Problems Governed by Linear Functional Differential Equations*

DOUGLAS C. REBER

Lefschetz Center for Dynamical Systems, Division of Applied Mathematics, Brown University, Providence, Rhode Island 02912

Received March 22, 1978

Aspects of the approximation and optimal control of systems governed by linear retarded nonautonomous functional differential equations (FDE) are considered. First, certain FDE are shown to be equivalent to corresponding abstract ordinary differential equations (ODE). Next, it is demonstrated that these abstract ODE may be approximated by difference equations in finite dimensional spaces. The optimal control problem for systems governed by FDE is then reduced to a sequence of mathematical programming problems. Finally, numerical results for two examples are presented and discussed.

1. INTRODUCTION

Our concern in this investigation is with the approximation and optimal control of systems governed by linear retarded nonautonomous functional differential equations (FDE). After presenting some basic properties of solutions of FDE in Section 2, we demonstrate in Section 3 that certain FDE are equivalent to corresponding abstract ordinary differential equations (ODE). This equivalence leads to two significant results. The first is the validity of a "variation of constants" representation of solutions in the state space $R^n \times L_2(-r, 0; R^n)$. (A similar result was obtained by Delfour [7]; such representations in the state space $C(-r, 0; R^n)$ are well known—see Hale [10, p. 207].) For our purposes, the importance of this observation lies in the fact that a compactness property of the variation of constants representation also obtains for the solution map of FDE. The second result is that a finite difference technique, similar to those used in the field of partial differential equations, may be employed to approximate solutions of FDE. These results are discussed in Sections 4 through 8.

* This research was supported, in part, by the National Science Foundation under Grant No. NSF-MCS76-07247, in part by the United States Air Force under Grant No. AF-AFOSR 76-3092, and in part while the author was a Corinna Borden Keen Research Fellow. These results were based on the author's dissertation at Brown University.

0022-0396/79/050193–40$02.00/0

The operator theoretic framework for the approximation of solutions of FDE requires an investigation of sufficient conditions (known as the stability and consistency conditions) for convergence of approximate solutions to the true solution. Other, more routine details of the particular scheme we have chosen are incorporated in the definitions and interrelationships of various spaces. The scheme itself has been studied by Delfour [7] by a more direct approach. One objective, therefore, of this investigation is the reformulation of an existing technique in such a manner that certain essential features are emphasized. An immediate additional benefit is that more general optimization problems than those considered by Delfour [7] (linear-quadratic) are seen to be easily handled.

The finite difference technique leads naturally to the definition of a sequence of mathematical programming problems. The original optimization problem (i.e. that which is governed by a linear FDE) is shown in Sections 9 and 10 to be the "limit" of these approximating problems, in the sense that the corresponding optimal controls, payoffs and trajectories all converge.

We then discuss numerical results for two examples in Sections 11 through 13. Standard techniques of numerical analysis were applied in each case to solve the approximating problems. The first example was chosen for its simplicity, so that an analytical solution would be readily available; the second is associated with a biochemical process.

Finally, some concluding remarks on the above technique are made in Section 14.

Most of the notation employed is standard. In particular, given $p \geqslant 1$, a closed interval I and a Banach space X, the symbol $L_p(I; X)$ will denote the set of (equivalence classes of) strongly Lebesgue measurable functions $f: I \to X$ for which $\int_I |f|^p < \infty$. $L_p(I; X)$ is made into a Banach space by definition of the usual norm $|\cdot|_{L_p}$. The Banach space of continuous functions with the supremum norm will be denoted by $C(I; X)$. $W_2^{(1)}(I; X)$ denotes the set of absolutely continuous functions from I to X whose derivatives are in $L_2(I; X)$. For Banach spaces X, Y the symbols $\mathscr{B}(X, Y)$, $\mathscr{B}(X)$ will represent the usual sets of continuous linear transformations with the uniform operator topology. The spaces R^n and $R^{n\times n}$ will be endowed with the euclidean and spectral norms, respectively.

Given arbitrary sets G, H with $G \subset H$, define $\mathrm{ch}(G, H)$ as the characteristic function of G. We shall assume that the positive integers m, n, κ and ν, the positive number r and the numbers a, b with $b > a$ are fixed. For a function f defined on $[a - r, b]$ and $t \in [a, b]$, the function f_t will be defined on $[-r, 0]$ by $f_t(\theta) = f(t + \theta)$. The symbol Δ will denote the set $\{(t, s): a \leqslant s \leqslant t \leqslant b\}$.

Several results are stated without proof in the sequel; unless otherwise indicated, proofs may be found in Reber [15].

The author wishes to express his appreciation for the helpful comments and suggestions of Professor H. T. Banks.

2. The Initial Value Problem

For the moment, consider an FDE as an equation which relates the derivative of a function $x\colon [a-r, b] \to R^n$ at time $t \in [a, b]$ to the values of a given function f and an operator $L(t, \cdot)$ acting on x_t as an element of L_2. Thus we write

$$\dot{x}(t) = L(t, x_t) + f(t)$$
$$(x(a), x_a) = (\eta, \phi).$$

It might appear that some simple FDE may not be formulated in this way. For example, consider $\dot{x}(t) = e(t)\, x(t-1)$ with $e \in L_2(a, b; R)$. The difficulty lies in the fact that for a given $(t, \phi) \in [a, b] \times L_2$ the value $e(t)\,\phi(t-1)$ is not well defined. This is merely an inconvenience, since we may reformulate the problem as

$$x(t) = \eta + \int_a^t [L(s, x_s) + f(s)]\, ds,$$
$$x_a = \phi$$

just as is done for ODE satisfying the Carathéodory conditions.

Several preliminary definitions are required to describe the FDE under consideration. Let the Banach spaces Z, Λ, E and F be given by

$$Z = R^n \times L_2(-r, 0; R^n)$$
$$\Lambda = L_2\left(a, b; \left(\bigtimes_0^{\nu} R^{n\times n}\right) \times L_2(-r, 0; R^{n\times n})\right),$$
$$E = C(a, b; Z) \qquad \text{and}$$
$$F = L_2(a, b; R^n).$$

A generic element of Z will be denoted by $\zeta = (\eta, \phi)$. The norms of Z and $(\bigtimes_0^{\nu} R^{n\times n}) \times L_2(-r, 0; R^{n\times n})$ are the usual norms for product spaces. A generic element of Λ will be denoted by $\lambda = (A_0, \ldots, A_\nu, D)$; the components $A_0, \ldots, A_\nu$ may be considered elements of $L_2(a, b; R^{n\times n})$ and the component D may be considered an element of $L_2([a, b] \times [-r, 0]; R^{n\times n})$. Let W denote the subset

$$W = \{(\eta, \phi) \in Z : \phi \in W_2^{(1)}, \eta = \phi(0)\}.$$

Now define the operator $L\colon [a, b] \times L_2(-r, 0; R^n) \times \Lambda \to R^n$ by

$$L(t, \phi) = L(t, \phi, \lambda) = \sum_0^{\nu} A_j(t)\, \phi(-\tau_j) + \int_{-r}^{0} D(t, \theta)\, \phi(\theta)\, d\theta$$

where $0 = \tau_0 < \cdots < \tau_\nu = r$ and $\lambda = (A_0, ..., A_\nu, D)$. We remark that although in a strict sense L is not well defined (point evaluations of λ and ϕ are required), no problem is encountered because these terms will appear under an integral sign in our usage below.

Thus we consider FDE of the form

$$\dot{x}(t) = L(t, x_t, \lambda) + f(t) \tag{2.1}$$

with $(\lambda, f) \in \Lambda \times F$ and initial value

$$(x(s), x_s) = \zeta \in Z, \qquad s \in [a, b] \text{ fixed.} \tag{2.2}$$

A function x is a *solution* of the initial value problem (IVP) given in Eqs. (2.1), (2.2) if: $x \in W_2^{(1)}(s, b; R^n)$, x satisfies Eq. (2.1) almost everywhere in $[s, b]$ and $(x(s), x_s)$ satisfies Eq. (2.2). (For convenience, when $s = b$ define a solution of the IVP in the obvious way.) Our discussion will concern only this restricted class of FDE.

We now describe some properties of solutions of Eqs. (2.1), (2.2).

2.1. THEOREM. *There is a unique solution of Eqs.* (2.1), (2.2) *which depends continuously on* $\gamma = (s, \zeta, \lambda, f)$ *in the sense that the map* $(t, \gamma) \to (x(t; \gamma), x_t(\gamma))$ *is continuous on* $\Delta \times Z \times \Lambda \times F$ *into* Z.

Define the operator $V: \Delta \times \Lambda \times Z \to Z$ by $V(t, s, \lambda)\zeta = (x(t; s, \zeta, \lambda, 0), x_t(s, \zeta, \lambda, 0))$. One may easily see from the above that V is continuous and is linear in ζ for fixed (t, s, λ). Therefore, whenever $B \subset \Lambda$ is relatively compact, there is, by the uniform boundedness principle, a constant M (depending on B) such that $\sup\{| V(t, s, \lambda)|_{\mathscr{B}(Z)}: (t, s, \lambda) \in \Delta \times B\} \leqslant M$. Uniqueness of solutions of the IVP implies that $V(t, s) = V(t, \tau)\, V(\tau, s)$ and $V(s, s) = I$ for all $a \leqslant s \leqslant \tau \leqslant t \leqslant b$. It is easily verified that $V(t, s, \lambda)$ takes W into itself for all $(t, s, \lambda) \in \Delta \times \Lambda$.

When dealing with linear ODE in R^n, the variation of constants formula provides a very useful explicit representation of solutions. The map Ψ defined below retains the form of this expression; that it indeed represents a solution of FDE, and in what sense, will be demonstrated in the next section.

Let $\Psi: Z \times \Lambda \times F \to E$ be given by

$$\Psi(\zeta, \lambda, f)(t) = V(t, a, \lambda)\, \zeta + \int_a^t V(t, s, \lambda)(f(s), 0)\, ds.$$

The existence of the integral in Z is assured by the uniform boundedness of V over $\Delta \times \{\lambda\}$ and the strong measurability of the map $s \to V(t, s, \lambda)(f(s), 0)$.

Now define $z: Z \times \Lambda \times F \to E$

$$z(\zeta, \lambda, f)(t) = (x(t, a, \zeta, \lambda, f), x_t(a, \zeta, \lambda, f)).$$

We list some properties of the functions z and Ψ which will be of interest later (e.g. in the proof of Theorem 3.14).

2.2. LEMMA. *The functions z and Ψ from $Z \times \Lambda \times F$ to E are continuous. Furthermore, for each fixed (ζ, λ) the map $f \to \Psi(\zeta, \lambda, f)$ is affine and compact.*

3. EQUIVALENCE OF FDE AND ABSTRACT ODE

In this section we shall establish a relationship between solutions of certain FDE and solutions of corresponding abstract ODE in the space Z. For this part of our discussion we will require that λ be contained in a proper subset of Λ. Therefore, define the set

$$\Lambda_c = C_1\left(a, b; \left(\bigtimes_0^{\nu} R^{n\times n}\right) \times L_2(-r, 0; R^{n\times n})\right)$$

of continuously differentiable functions on $[a, b]$. Define a norm on Λ_c by $|\lambda|_c = \sum_0^{\nu} |A_j|_c + |D|_c$, where each A_j is considered an element of $C(a, b; R^{n\times n})$ and D is considered an element of $C(a, b; L_2(-r, 0; R^{n\times n}))$.

It is not difficult to see [8, Lemma 19, p. 298] that for a separable Banach space X, $C_1(a, b; X)$ is dense in $L_2(a, b; X)$. Thus Λ_c is dense in Λ.

The following lemma is given for reference; its proof follows from Theorem 2.1.

3.1. LEMMA. *Suppose $\gamma = (\zeta, \lambda, f) \in W \times \Lambda \times F$. Then the function $(t, s) \to x_t(s, \gamma) \in C(-r, 0; R^n)$ is continuous on Δ. Moreover, the function $(t, s) \to L(t, x_t(s, \gamma))$ is continuous on Δ for $\gamma \in W \times \Lambda_c \times F$.*

Let $\mathscr{A}$: $[a, b] \times \Lambda_c \times W \to Z$ be given by $\mathscr{A}(t)(\psi(0), \psi) = \mathscr{A}(t; \lambda)(\psi(0), \psi) = (L(t, \psi, \lambda), \dot{\psi})$. Consider the initial value problem

$$\dot{y}(t) = \mathscr{A}(t; \lambda)\, y(t) \qquad t \in [s, b], \quad s \in [a, b] \text{ fixed,} \tag{3.1}$$

$$y(s) = \zeta \in W. \tag{3.2}$$

A *solution* of Eqs. (3.1), (3.2) on $[s, b]$ is a continuous function y: $[s, b] \to W$ which satisfies Eq. (3.1) on $[s, b]$ (one-sided derivatives to be taken at the endpoints) and is such that $y(s) = \zeta$. (For convenience, when $s = b$ define a solution of the IVP in the obvious way.)

The following definition is essentially that which appears in Krein [11, p. 193].

3.2. Definition. The IVP given ir Eqs. (3.1), (3.2) is *uniformly correct* if:

(i) for each $s \in [a, b]$ and $\zeta \in W$ there is a unique solution of the IVP,

(ii) each solution $y(t, s, \zeta, \lambda)$ and its derivative $(d/dt)\, y(t, s, \zeta, \lambda)$ are continuous for $(t, s) \in \Delta$ and fixed (ζ, λ), and

(iii) the solution depends continuously on the initial data in the sense that if $\zeta_i \in W$ and $\zeta_i \to 0$ then the corresponding solutions converge to zero uniformly relative to $(t, s) \in \Delta$.

The results given in Lemma 3.3, Lemma 3.4 and Theorem 3.5 rely on the fact that certain operators are maximal dissipative. (The definition of maximal dissipative operators and some standard results concerning them may be found in Krein [11, p. 86f.].) The idea of redefining the inner product in the proof of Lemma 3.3 is related to similar definitions in [3] and [18].

3.3. Lemma. *Given $\lambda \in \Lambda_c$, there is an equivalent inner product topology on Z and a positive $\omega = \omega(\lambda)$ such that $\mathscr{A}(t; \lambda) - \omega(\lambda) I$ is maximal dissipative for each $t \in [a, b]$.*

Proof. Define $Z_\lambda = R^n \times L_2(-r, 0; R^n)$ with inner product $\langle \cdot, \cdot \rangle_\lambda$, where

$$\langle (\eta, \phi), (\alpha, \psi) \rangle_\lambda = \langle \eta, \alpha \rangle_{R^n} + \sum_1^\nu \beta_j \int_{-\tau_j}^{-\tau_{j-1}} \langle \phi(\theta), \psi(\theta) \rangle_{R^n}\, d\theta$$

and $\beta_j = 1 + \sum_{i=j}^\nu |A_i|_c$. It is easy to see that

$$|(\eta, \phi)|_Z \leqslant |(\eta, \phi)|_\lambda \leqslant \beta_1^{1/2} |(\eta, \phi)|_Z,$$

so the topologies are equivalent.

If $(\eta_j, \phi_j) \to (\eta, \phi)$ and $\mathscr{A}(t)(\eta_j, \phi_j) \to (\alpha, \psi)$ then:

(i) $\phi_j \to \phi$ and $\dot{\phi}_j \to \psi$ in L_2 imply that $\phi(\theta) = \phi(0) + \int_0^\theta \psi(\sigma)\, d\sigma$ and $\phi_j \to \phi$ in $C(-r, 0; R^n)$, and

(ii) $\eta_j \to \eta$ in R^n implies that $\phi(0) = \eta$.

So $(\eta, \phi) \in W$. Furthermore $\mathscr{A}(t)(\eta_j, \phi_j) \to \mathscr{A}(t)(\eta, \phi)$, i.e. $\mathscr{A}(t)$ is closed.

Let $\omega(\lambda) = 1 + |\lambda|_c$. To see that $\mathscr{A}(t) - \omega I$ is dissipative on its domain W:

$$\begin{aligned} &\langle \mathscr{A}(t)(\eta, \phi), (\eta, \phi) \rangle_\lambda \\ &\quad = \sum_0^\nu \langle \eta, A_j(t)\, \phi(-\tau_j) \rangle_{R^n} + \left\langle \eta, \int_{-r}^0 D(t, \theta)\, \phi(\theta)\, d\theta \right\rangle_{R^n} \\ &\qquad + \sum_1^\nu \beta_j \int_{-\tau_j}^{-\tau_{j-1}} \langle \dot{\phi}(\theta), \phi(\theta) \rangle_{R^n}\, d\theta \end{aligned}$$

$$
\begin{aligned}
&\leqslant \left[| A_0(t)| + (1/2) \sum_1^{\nu} | A_j(t)| \right] | \eta |^2 + (1/2) \sum_1^{\nu} | A_j(t)| \, | \phi(-\tau_j)|^2 \\
&\quad + (1/2) \, | D(t)| \, | \eta |^2 + (1/2) \, | D(t)| \, | \phi |^2 + (1/2) \sum_1^{\nu} | A_j |_c \, | \eta |^2 \\
&\quad - (1/2) \sum_1^{\nu} | A_j |_c \, | \phi(-\tau_j)|^2 + (1/2) \, | \eta |^2 - (1/2) \, | \phi(-r)|^2 \\
&\leqslant \left[(1/2)(1 + | D |_c) + \sum_0^{\nu} | A_j |_c \right] | \eta |^2 + (1/2) \, | D |_c \, | \phi |^2 \\
&\leqslant (| \lambda |_c + 1/2) \, |(\eta, \phi)|_\lambda^2 \\
&= (\omega - 1/2) \, |(\eta, \phi)|_\lambda^2 ,
\end{aligned}
$$

where we have repeatedly used the inequality $cd \leqslant (1/2)(c^2 + d^2)$.

It remains to show that $\mathscr{A}(t) - \omega I$ is onto, which will establish maximal dissipativeness. Fix (α, ψ) in Z. Let $\phi \in L_2$ be a solution of the ODE $\dot{\phi} - \omega\phi = \psi$; in particular let $\phi(\theta) = e^{\omega\theta}\phi(0) + \int_0^\theta e^{\omega(\theta-s)}\psi(s)\, ds$. Clearly $(\phi(0), \phi) \in W$; we need only demonstrate that $\phi(0) \in R^n$ may be chosen so that $L(t, \phi) - \omega\phi(0) = \alpha$. Write

$$
L(t, \phi, \lambda) = \sum_0^{\nu} A_j(t)\, \phi(-\tau_j) + \int_{-r}^0 D(t, \theta)\, \phi(\theta)\, d\theta
$$

as $H(t)\, \phi(0) + G(t, \psi)$, where

$$
H(t) = \sum_0^{\nu} e^{-\omega\tau_j} A_j(t) + \int_{-r}^0 e^{\omega\theta} D(t, \theta)\, d\theta,
$$

and

$$
G(t, \psi) = \sum_0^{\nu} A_j(t) \int_0^{-\tau_j} e^{\omega(-\tau_j - s)} \psi(s)\, ds + \int_{-r}^0 D(t, \theta) \int_0^\theta e^{\omega(\theta-s)} \psi(s)\, ds\, d\theta.
$$

The equation $L(t, \phi) - \omega\phi(0) = \alpha$ has a solution if $H(t) - \omega I$ is invertible, i.e. if $(H(t) - \omega I)\, \phi(0) = \alpha - G(t, \psi)$ has a solution. Observe that

$$
\begin{aligned}
| H(t)| &\leqslant \sum_0^{\nu} | A_j |_c + \left\{ \int_{-r}^0 e^{2\omega\theta}\, d\theta \right\}^{1/2} | D(t)|_{L_2} \\
&\leqslant | \lambda |_c
\end{aligned}
$$

since $(\int_{-r}^0 e^{2\omega\theta}\, d\theta)^{1/2} = [(1/2\omega)(1 - e^{-2\omega r})]^{1/2} \leqslant 1$ by choice of $\omega = 1 + | \lambda |_c \geqslant 1$. Consequently $(H(t) - \omega I)^{-1}$ exists and has norm no greater than one. Hence $\mathscr{A}(t) - \omega I$ is onto. ∎

3.4. LEMMA. $|(\mathscr{A}(t) - \mu I)^{-1}|_{\mathscr{B}(Z_\lambda)} \leqslant (\mu - \omega + 1/2)^{-1}$ *for all* $\mu \geqslant \omega(\lambda)$.

Proof. Suppose $\zeta \in W$ and $\mu \geqslant \omega$. Let $\xi = (\mathscr{A}(t) - \mu I)\zeta$. Then

$$-\langle \xi, \zeta\rangle_\lambda = -\langle \mathscr{A}(t)\zeta, \zeta\rangle_\lambda + \mu\langle \zeta, \zeta\rangle_\lambda ,$$

so $-\langle \xi, \zeta\rangle_\lambda \geqslant (\mu - \omega + 1/2)\langle \zeta, \zeta\rangle_\lambda$. Since $-\langle \xi, \zeta\rangle_\lambda \leqslant |\xi|_\lambda |\zeta|_\lambda$ we have

$$|\zeta|_\lambda \leqslant (\mu - \omega + 1/2)^{-1} |\xi|_\lambda = (\mu - \omega + 1/2)^{-1} |(\mathscr{A}(t) - \mu I)\zeta|_\lambda .$$

Therefore $|(\mathscr{A}(t) - \mu I)^{-1}|_{\mathscr{B}(Z_\lambda)} \leqslant (\mu - \omega + 1/2)^{-1}$. ∎

The following theorem will enable us to conclude that solutions of the IVP (3.1), (3.2) (if they exist) are unique and depend continuously on the initial data in the sense of part (iii) of Definition 3.2. The basic idea of the proof may be found in Krein [11, p. 204].

3.5. THEOREM. *Every solution* y *of Eqs.* (3.1), (3.2) *satisfies the inequality*

$$|y(t)|_Z \leqslant M(\lambda)\, e^{\omega(\lambda)(t-s)} |\zeta|_Z ,$$

where $M(\lambda) = (1 + \sum_1^\nu |A_j|_c)^{1/2}$.

Proof. We again find it convenient to use $\langle \cdot, \cdot\rangle_\lambda$. By the definitions of derivative in Z and of a solution of the IVP,

$$(y(t + \epsilon, s) - y(t, s))/\epsilon \to \mathscr{A}(t)\, y(t)$$

as $\epsilon \to 0$. Hence

$$\begin{aligned} y(t + \epsilon, s) &= [I + \epsilon\mathscr{A}(t)]\, y(t, s) + o(\epsilon) \\ &= [I - \epsilon^2\mathscr{A}^2(t)][I - \epsilon\mathscr{A}(t)]^{-1}\, y(t, s) + o(\epsilon) \\ &= [I - \epsilon\mathscr{A}(t)]^{-1}\, y(t, s) - \epsilon^2\mathscr{A}(t)[I - \epsilon\mathscr{A}(t)]^{-1}\,\mathscr{A}(t)\, y(t, s) + o(\epsilon). \end{aligned}$$

In view of Lemma 3.4, for $1/\epsilon > \omega(\lambda)$ we have

$$\begin{aligned} |[I - \epsilon\mathscr{A}(t)]^{-1}|_{\mathscr{B}(Z_\lambda)} &= (1/\epsilon)\, |[(1/\epsilon)\, I - \mathscr{A}(t)]^{-1}|_{\mathscr{B}(Z_\lambda)} \\ &\leqslant (1/\epsilon)[(1/\epsilon) - \omega]^{-1} \\ &= 1 + \epsilon\omega + o(\epsilon). \end{aligned}$$

Thus

$$|y(t + \epsilon, s)|_\lambda \leqslant |y(t, s)|_\lambda(1 + \epsilon\omega) + \epsilon\, |\mathscr{A}(t)[(1/\epsilon)I - \mathscr{A}(t)]^{-1}\mathscr{A}(t)\, y(t, s)|_\lambda + o(\epsilon).$$

Since $\mathscr{A}(t) - \omega I$ is maximal dissipative, we may apply a result of Pazy [14, Lemma 3.2, p. 10] to conclude that for all $w \in Z$,

$$(1/\epsilon)([\omega + (1/\epsilon)]I - \mathscr{A}(t))^{-1} w \to w \qquad \text{as} \quad \epsilon \downarrow 0.$$

This implies that $\mu[\mu I - \mathscr{A}(t)]^{-1}w \to w$ as $\mu \to \infty$ (multiply both sides by $1 + \epsilon\omega$; let $\mu = (1/\epsilon) + \omega$). For $1/\epsilon > \omega$ we have

$$(1/\epsilon)[(1/\epsilon)I - \mathscr{A}(t)]^{-1}w - w = \mathscr{A}(t)[(1/\epsilon)I - \mathscr{A}(t)]^{-1}w.$$

Therefore $\mathscr{A}(t)[(1/\epsilon)I - \mathscr{A}(t)]^{-1}w \to 0$ as $\epsilon \downarrow 0$ for all $w \in Z$, and so

$$(| y(t+\epsilon, s)|_\lambda - | y(t, s)|_\lambda)/\epsilon \leqslant \omega \,| y(t, s)|_\lambda + O(\epsilon) + o(\epsilon)/\epsilon$$

for all ϵ with $1/\epsilon > \omega$. Hence

$$\bar{D}^+ \,| y(t, s)|_\lambda \leqslant \omega \,| y(t, s)|_\lambda$$

where $\bar{D}^+$ indicates the upper right derivative.

If $f \in C(s, t; R)$, then (see [19, pp. 239, 240]) $\bar{D}^+ f(\theta) \geqslant 0$ on $[s, t]$ implies that $f(t) \geqslant f(s)$. In particular, for $f(\theta) = -e^{-\omega(\lambda)\theta} \,| y(\theta, s)|_\lambda$ we find that

$$| y(t, s)|_\lambda \leqslant e^{\omega(\lambda)(t-s)} \,| y(s, s)|_\lambda , \qquad \text{i.e.}$$
$$| y(t, s)|_Z \leqslant M(\lambda)\, e^{\omega(\lambda)(t-s)} \,| \zeta |_Z . \quad \blacksquare$$

Several more lemmas are required to establish the relationship between FDE and abstract ODE. The proof of the following result follows a standard argument (see [19, p. 239]); we state it here for sake of completeness.

3.6. Lemma. *Suppose X is a Banach space. Let $x \in C(a, b; X)$ be such that $\dot{x}^+(t)$ exists and is continuous on $[a, b)$. Then $\dot{x}(t)$ exists on $[a, b]$ (one-sided derivatives to be taken at a, b) and*

$$x(t) = x(a) + \int_a^t \dot{x}(s)\, ds = x(a) + \int_a^t \dot{x}^+(s)\, ds$$

for all $t \in [a, b]$.

3.7. Lemma. (1) *For every $\lambda \in \Lambda_c$ and $s \in [a, b)$ the derivative $\partial/\partial t V(t, s)\zeta$ exists in Z for $t \in [s, b]$ and equals $\mathscr{A}(t)\, V(t, s)\zeta$, if and only if $\zeta \in W$.*

(2) *Similarly, if $t \in (a, b]$ and $\zeta \in W$, the derivative $\partial/\partial s V(t, s)\zeta$ exists in Z for $s \in [a, t]$ and equals $-V(t, s)\, \mathscr{A}(s)\zeta$. (One-sided derivatives are to be taken at the endpoints.)*

Proof. (1) We first show that for $s \in [a, b)$, $\partial^+/\partial t\, V(t, s)\zeta \,|_{t=s}$ exists if $\zeta \in W$.

Suppose $(\eta, \phi) = \zeta \in W$. Then Lemma 3.1 implies that

$$(1/\epsilon) \int_s^{s+\epsilon} L(\theta, x_\theta)\, d\theta \to L(s, x_s)$$

as $\epsilon \downarrow 0$. It is an elementary fact (see [9, p. 254]) that $\dot{x} \in L_2(s - r, b)$ implies $(1/\epsilon)(x_{s+\epsilon} - x_s) \to (\dot{x})_s$ in $L_2(-r, 0)$ as $\epsilon \downarrow 0$.

Now suppose that the right derivative $\partial^+/\partial t\, V(t, s)\zeta\,|_{t=s}$ exists. Then there is a point $(\alpha, \psi) \in Z$ such that $(1/\epsilon)(x_{s+\epsilon} - x_s) \to \psi$ in L_2 and $(1/\epsilon)[x(s+\epsilon) - x(s)] \to \alpha$ in R^n as $\epsilon \downarrow 0$. Let $v(\cdot; \epsilon)$: $[-r, 0] \to R^n$ be given by

$$v(\theta; \epsilon) = (1/\epsilon) \int_{s+\theta}^{s+\theta+\epsilon} x(\sigma)\, d\sigma.$$

Then the maps $\theta \to d/d\theta\, v(\theta; \epsilon)$ and $\theta \to (1/\epsilon)[x(s+\theta+\epsilon) - x(s+\theta)]$ are equal in L_2. Observe that $v(\cdot; \epsilon) \to x_s = \phi$ and $d/d\theta\, v(\cdot; \epsilon) \to \psi$ in L_2 as $\epsilon \downarrow 0$. By the continuity of x on $[s, b]$, $v(0; \epsilon) \to x(s) = \eta$.

Let ξ: $[s - r, b] \to R^n$ be given by

$$\xi(t) = \begin{cases} \eta + \displaystyle\int_0^{t-s} \psi(\theta)\, d\theta & t \in [s-r, s] \\ x(t) & t \in [s, b] \end{cases}$$

Clearly $d/d\theta\, v(\cdot; \epsilon) \to \dot{\xi}_s = \psi$ in L_2 as $\epsilon \downarrow 0$. Since $v(0; \epsilon) \to \eta = \xi(s)$, we have $v(\cdot; \epsilon) \to \xi_s$ in L_2 as $\epsilon \downarrow 0$, which implies that $\xi_s = \phi$ in L_2. Consequently ξ is an absolutely continuous representor for x on $[s - r, b]$. Thus $x \in W_2^{(1)}(s - r, b)$, so $\alpha = \lim_{\epsilon\downarrow 0}(1/\epsilon) \int_s^{s+\epsilon} L(\theta, x_\theta)\, d\theta = L(s, x_s) = L(s, \phi)$. Therefore $(\eta, \phi) \in W$ and the right derivative equals $\mathscr{A}(s)(\eta, \phi)$.

Fix $s \in [a, b)$. For $t \in [s, b)$, $\epsilon > 0$ and $\zeta \in W$ we have

$$(1/\epsilon)[V(t+\epsilon, s) - V(t, s)]\zeta = (1/\epsilon)[V(t+\epsilon, t) - I]\, V(t, s)\zeta.$$

The limit of the right-hand side as $\epsilon \downarrow 0$ is $\mathscr{A}(t)\, V(t, s)\zeta$.

Using the fact that $\zeta \in W$, the definitions of $\mathscr{A}(t)$ and $V(t, s)$, and Lemma 3.1, we may conclude that the function $t \to \mathscr{A}(t)\, V(t, s)\zeta$ is continuous on $[s, b)$. Hence the function $t \to V(t, s)\zeta$ is continuously right differentiable on $[s, b)$. Therefore, Lemma 3.6 implies that $V(t, s)\zeta = \zeta + \int_s^t \mathscr{A}(\theta)\, V(\theta, s)\zeta\, d\theta$ for all $t \in [s, b]$, and thus the desired conclusion.

(2) Fix $t \in (a, b]$. Then for $s \in [a, t)$, $\epsilon > 0$ sufficiently small, and $\zeta \in W$ we have

$$(1/\epsilon)[V(t, s+\epsilon) - V(t, s)]\zeta = -V(t, s+\epsilon)(1/\epsilon)[V(s+\epsilon, s) - I]\zeta.$$

The limit of the right-hand side as $\epsilon \downarrow 0$ is $-V(t, s)\, \mathscr{A}(s)\zeta$ (by the results of part (1), the uniform boundedness of $V(t, s)$ over Δ and the continuity of $s \to V(t, s)w$ for all $w \in Z$).

The strong continuity of $\mathscr{A}(s)$ over W and the uniform boundedness of $V(t, s)$ over Δ imply that the function $s \to -V(t, s)\, \mathscr{A}(s)\zeta$ is continuous. Therefore, we may again use Lemma 3.6 to conclude that for $s \in [a, t]$,

$$V(t, s)\, \zeta = \zeta + \int_s^t V(t, \sigma)\, \mathscr{A}(\sigma)\, \zeta\, d\sigma. \quad \blacksquare$$

We may now relate the solutions of Eqs. (3.1), (3.2) to those of Eqs. (2.1), (2.2) when $f = 0$.

3.8. THEOREM. *For* $\lambda \in \Lambda_c$ *and* $\zeta \in W$, $V(\cdot, s, \lambda)\zeta$ *is the unique solution of Eqs.* (3.1), (3.2). *Furthermore, this* IVP *is uniformly correct.*

Proof. Theorem 3.5 and Lemma 3.7 imply that $V(\cdot, s, \lambda)\zeta$ is the unique solution of Eqs. (3.1), (3.2). Uniform correctness then follows immediately from Lemma 3.1 and the corresponding properties of solutions of Eqs. (2.1), (2.2). ∎

We turn now to the inhomogeneous case. Thus we consider the IVP

$$\dot{y}(t) = \mathscr{A}(t)\, y(t) + g(t), \tag{3.3}$$

$$y(s) = \zeta \in W, \qquad s \in [a, b] \text{ fixed.} \tag{3.4}$$

While Eqs. (3.3), (3.4) represent the problem we wish to solve, technical considerations (to be discussed below) require some slight modifications. In particular, for s as in Eq. (3.4) and $\omega = \omega(\lambda)$, define $\mathscr{A}_\omega(t) = \mathscr{A}(t) - \omega I$ and $g_\omega(t) = e^{\omega(s-t)} g(t)$. We shall consider the IVP given by Eq. (3.4) and

$$\dot{y}(t) = \mathscr{A}_\omega(t)\, y(t) + g_\omega(t). \tag{3.5}$$

For operators $U(t, s)\colon Z \to Z$ defined for $(t, s) \in \Delta$ we list properties that will be discussed in the sequel (see [11, p. 195]):

1°. the operator $U(t, s)$ is bounded on Z uniformly relative to $(t, s) \in \Delta$;

2°. the operator $U(t, s)$ is strongly continuous in (t, s) on Δ;

3°. $U(t, s) = U(t, \tau)\, U(\tau, s)$, $U(s, s) = I$ for $a \leqslant s \leqslant \tau \leqslant t \leqslant b$;

4°. (i) the operator $U(t, s)$ maps W into itself,

(ii) the operator $\mathscr{A}_\omega(t)\, U(t, s)\, \mathscr{A}_\omega^{-1}(s)$ is bounded and strongly continuous on Z for $(t, s) \in \Delta$;

5°. on W the operator $U(t, s)$ is strongly differentiable relative to t and s, with $\partial/\partial t\, U(t, s)\zeta = \mathscr{A}_\omega(t)\, U(t, s)\zeta$ and $\partial/\partial s\, U(t, s)\zeta = -U(t, s)\, \mathscr{A}_\omega(s)\zeta$.

In addition to the hypotheses listed in Lemma 3.9 below, there is a standing assumption in Krein [11, p. 195] under which this lemma is proved. In particular he assumes (as we also do) that $\mathscr{A}_\omega(t)$ has a bounded inverse satisfying $\sup\{|\mathscr{A}_\omega(a)\, \mathscr{A}_\omega^{-1}(s)| : s \in [a, b]\} < \infty$. This requirement is treated in Lemmas 3.10 and 3.11. The proof of Lemma 3.9 may be found in Krein [11, Theorem 3.3, p. 197].

3.9. LEMMA. *If, in addition to the above assumption, we have:*

(a) *the* IVP *given by Eq.* (3.4) *and the homogeneous part of Eq.* (3.5) *is uniformly correct*;

(b) $\mathscr{A}_\omega(t)$ *is strongly continuously differentiable on* W *for* $t \in [a, b]$;

(c) $U(t, s)$ *satisfies properties* 1°–5°; *and*

(d) $g_\omega(t) = e^{\omega(a-t)} g(t)$ *is continuously differentiable,*

then $y(t) = \int_a^t U(t, s)\, g_\omega(s)\, ds$ *yields a solution of Eq.* (3.5) *with initial value zero at* $t = a$.

Clearly $e^{\omega(s-t)} y(t)$ is a solution of Eqs. (3.4), (3.5) if and only if $y(t)$ is a solution of Eqs. (3.3), (3.4). Therefore, the homogeneous part of Eq. (3.5) is uniformly correct, and its solution operator is given by $V_\omega(t, s) = e^{\omega(s-t)} V(t, s)$.

From our previous results, it is not difficult to see that $V_\omega(t, s)$ satisfies properties 1°–3°, 4°(i), 5° and that $\mathscr{A}_\omega(t)$ is strongly continuously differentiable on W for $t \in [a, b]$ whenever $\lambda \in \Lambda_c$.

3.10. LEMMA (cf. Krein [11, pp. 176, 177]). *Condition* 4°(ii) *obtains if zero is a regular point of* $\mathscr{A}_\omega(t)$ *for all* $t \in [a, b]$ *and* $K = \sup\{| \mathscr{A}_\omega(a)\, \mathscr{A}_\omega^{-1}(t)|_{\mathscr{B}(Z)}: a \leqslant t \leqslant b\}$ *is finite.*

Proof. Write $\mathscr{A}_\omega(t)\, V_\omega(t, s)\, \mathscr{A}_\omega^{-1}(s)$ as $[\mathscr{A}_\omega(t)\, V_\omega(t, s)\, \mathscr{A}_\omega^{-1}(a)][\mathscr{A}_\omega(a)\, \mathscr{A}_\omega^{-1}(s)]$. The map $(t, s) \to \mathscr{A}_\omega(t)\, V_\omega(t, s)$ is strongly continuous on W (this may be established as in the proof of Lemma 3.7, part (1)). Hence $(t, s) \to \mathscr{A}_\omega(t)\, V_\omega(t, s)\, \mathscr{A}_\omega^{-1}(a)$ is strongly continuous on Z. For each fixed (t, s), $\mathscr{A}_\omega(t)\, V_\omega(t, s)\, \mathscr{A}_\omega^{-1}(a)$ is a closed linear operator defined on all of Z; therefore it is bounded by the closed graph theorem. Strong continuity with respect to (t, s) and compactness of Δ imply by the Banach–Steinhaus theorem that the operators $\mathscr{A}_\omega(t)\, V_\omega(t, s)\, \mathscr{A}_\omega^{-1}(a)$ are uniformly bounded.

For $t, t + \sigma \in [a, b]$ and any $\zeta \in Z$ we have

$$\begin{aligned} &| \mathscr{A}_\omega(a)\, \mathscr{A}_\omega^{-1}(t + \sigma)\, \zeta - \mathscr{A}_\omega(a)\, \mathscr{A}_\omega^{-1}(t)\, \zeta | \\ &\quad = | \mathscr{A}_\omega(a)\, \mathscr{A}_\omega^{-1}(t + \sigma)[\mathscr{A}_\omega(t) - \mathscr{A}_\omega(t + \sigma)]\, \mathscr{A}_\omega^{-1}(t)\, \zeta | \\ &\quad \leqslant K\, |[\mathscr{A}_\omega(t) - \mathscr{A}_\omega(t + \sigma)]\, \mathscr{A}_\omega^{-1}(t)\, \zeta |. \end{aligned}$$

The right-hand side goes to zero as $\sigma \to 0$ by the strong continuity of $\mathscr{A}_\omega(t)$ on W. The conclusion follows immediately. ∎

We now show that the second hypothesis of Lemma 3.10 is satisfied for the operators $\mathscr{A}_\omega(t)$. (Lemma 3.3 shows that the first hypothesis is satisfied by $\mathscr{A}_\omega(t)$.) These operators were introduced because the operators $\mathscr{A}(t)$ do not necessarily satisfy a corresponding hypothesis.

3.11. LEMMA. *There is a constant K such that* $|\mathscr{A}_\omega(a)\mathscr{A}_\omega^{-1}(t)|_{\mathscr{B}(Z)} \leqslant K$ *for all* $t \in [a, b]$.

Proof. We may write

$$(\mathscr{A}(a) - \omega I)(\mathscr{A}(t) - \omega I)^{-1} = e_1 + e_2,$$

where $e_1 = [\mathscr{A}(a) - \mathscr{A}(t)](\mathscr{A}(t) - \omega I)^{-1}$ and $e_2 = (\mathscr{A}(t) - \omega I)(\mathscr{A}(t) - \omega I)^{-1} = I$. For $(\eta, \phi) \in W$,

$$|[\mathscr{A}(a) - \mathscr{A}(t)](\eta, \phi)|_Z = |L(a, \phi) - L(t, \phi)|_{R^n} \leqslant 2\rho\, |\lambda|_c\, |\phi|_c$$

where $\rho = \max\{1, r^{1/2}\}$. From the proof of Lemma 3.3 when $(\eta, \phi) = (\mathscr{A}(t) - \omega I)^{-1}(\alpha, \psi)$ we have $|\phi|_c \leqslant |\phi(0)| + r^{1/2}\, |\psi|_{L_2}$ and

$$|\eta| \leqslant |(H(t) - \omega I)^{-1}|\, |\alpha - G(t, \psi)| \leqslant |\alpha| + |G(t, \psi)|.$$

Observe that

$$|G(t, \psi)| \leqslant \sum_0^\nu |A_j|_c\, r^{1/2}\, |\psi|_{L_2} + |D|_c\, r\, |\psi|_{L_2}$$

$$\leqslant r^{1/2}\rho\, |\lambda|_c\, |\psi|_{L_2}.$$

Thus $|\eta| \leqslant |\alpha| + r^{1/2}\rho\, |\lambda|_c\, |\psi|_{L_2}$. Consequently

$$|\phi|_c \leqslant |\alpha| + r^{1/2}(1 + \rho\, |\lambda|_c)\, |\psi|_{L_2} \leqslant d\, |(\alpha, \psi)|_Z$$

where $d = 2[1 + r^{1/2}(\rho\, |\lambda|_c + 1)]$. So $|e_1|_{\mathscr{B}(Z)} \leqslant 2\, |\lambda|_c\, \rho d$; therefore $|\mathscr{A}_\omega(a)\mathscr{A}_\omega^{-1}(t)|_{\mathscr{B}(Z)} \leqslant 2\, |\lambda|_c\, \rho d + 1$. Let $K = 2\, |\lambda|_c\, \rho d + 1$. ∎

The following two lemmas are useful in establishing the validity of the term "variation of constants formula" as regards the function Ψ.

3.12. LEMMA. *For* $(\zeta, \lambda, f) \in W \times \Lambda_c \times C_1(a, b; R^n)$, $\Psi(\zeta, \lambda, f)$ *is the unique solution of Eqs.* (3.3), (3.4) *with* $g(t) = (f(t), 0)$ *and* $s = a$.

Proof. Solutions, if they exist, are unique because the equation is linear and its homogeneous part has unique solutions.

The preceding discussion allows us to conclude that

$$V_\omega(t, a)\, \zeta + \int_a^t V_\omega(t, \sigma)\, e^{\omega(a-\sigma)}(f(\sigma), 0)\, d\sigma$$

is the unique solution of Eq. (3.5) for $z(a) = \zeta$. Observe that this expression may be rewritten as

$$\begin{aligned}
& e^{\omega(a-t)} V(t, a)\, \zeta + \int_a^t e^{\omega(\sigma-t)} V(t, \sigma)\, e^{\omega(a-\sigma)}(f(\sigma), 0)\, d\sigma \\
&= e^{\omega(a-t)} \left\{ V(t, a)\, \zeta + \int_a^t V(t, \sigma)(f(\sigma), 0)\, d\sigma \right\} \\
&= e^{\omega(a-t)} \Psi(\zeta, \lambda, f)(t). \quad \blacksquare
\end{aligned}$$

3.13. Lemma. *For $(\zeta, \lambda, f) \in W \times \Lambda_c \times C(a, b; R^n)$, $z(\zeta, \lambda, f)$ is a solution of Eqs.* (3.3), (3.4) *with $g = (f, 0)$ and $s = a$.*

Proof. By Lemma 3.1 the function $s \to L(s, x_s)$ is continuous on $[a, b]$. Since f is also continuous,

$$\begin{aligned}
\lim_{\epsilon \to 0} (1/\epsilon)[x(t+\epsilon) - x(t)] &= \lim (1/\epsilon) \int_t^{t+\epsilon} [L(s, x_s) + f(s)]\, ds \\
&= L(t, x_t) + f(t).
\end{aligned}$$

Furthermore, $\dot{x} \in L_2(a - r, b)$ implies $\lim_{\epsilon \to 0}(1/\epsilon)[x_{t+\epsilon} - x_t] = (\dot{x})_t$ in L_2 for all $t \in [a, b]$. $\blacksquare$

We are now in a position to state the fundamental result of this section. Its proof follows immediately from Lemmas 2.2, 3.12 and 3.13 and the density of $W \times \Lambda_c \times C_1$ in $Z \times \Lambda \times F$. An important implication of this theorem is that the map $f \to z(\zeta, \lambda, f)$ is affine and compact for all fixed $(\zeta, \lambda) \in Z \times \Lambda$ (see Lemma 2.2). This fact will be used in Section 8.

3.14. Theorem. *The solution map $z(\zeta, \lambda, f)$ and the variation of constants representation $\Psi(\zeta, \lambda, f)$ are identical on $Z \times \Lambda \times F$.*

Theorem 3.15 below provides the basis for the approximation scheme to be discussed in Sections 4 through 8.

3.15. Theorem. *For $(\zeta, \lambda, f) \in W \times \Lambda_c \times F$, $z(\zeta, \lambda, f)$ is the unique solution of*

$$y(t) = \zeta + \int_a^t [\mathscr{A}(s)\, y(s) + (f(s), 0)]\, ds.$$

Proof. One may easily show that for $(\zeta, \lambda, f) \in W \times \Lambda_c \times F$ the function $\mathscr{A}(\cdot)\, z(\zeta, \lambda, f)(\cdot)$ is continuous on $[a, b]$. Hence the integral exists. Using the definitions of $\mathscr{A}$ and z, one may then verify that z is a solution of the integral equation. Uniqueness follows from Theorem 3.8. $\blacksquare$

In summary, we have found (Theorem 3.4) that the "variation of constants" representation, Ψ, is indeed the same as the solution map $z(t) = (x(t), x_t)$ of FDE. We have seen (Lemma 3.12) that for sufficiently smooth parameters

ζ, λ and f, Ψ (hence also z) is the solution of the corresponding abstract ODE. From this it followed (Theorem 3.15) that for arbitrary f in L_2, z was a solution of the above integral equation. We emphasize that this integral equation is equivalent to Eqs. (3.3), (3.4) only if f is continuous.

Finally, we remark that instead of using the FDE theory from Section 2 to show that the IVP of Eqs. (3.1), (3.2) is uniformly correct, we could have invoked a result of Krein [11, Theorem 3.11, p. 208] after our Lemma 3.4. This was not done for two reasons. The first is that some material required in the proof of Krein's theorem has not been discussed; the second reason is that properties of solutions of Eqs. (3.1), (3.2) imply (once one has proved uniqueness of solutions of Eqs. (2.1), (2.2)) corresponding results for the FDE only when $(\zeta, \lambda) \in W \times \Lambda_c$. Consequently the FDE theory had to be established independently in order to obtain results for general $(\zeta, \lambda) \in Z \times \Lambda$.

4. Factor Convergence of Solutions

We have shown (Theorem 3.15) that solving the equation

$$y(t) = \zeta + \int_a^t [\mathscr{A}(s; \lambda)\, y(s) + (f(s), 0)]\, ds \tag{4.1}$$

is in many cases equivalent to solving Eqs. (2.1), (2.2). In this chapter we present a finite difference method which leads to a very convenient approximation scheme for Eq. (4.1); this method may easily be implemented on a computer.

Finite difference schemes are based on the concept of factor convergence; a good introduction to this idea may be found in Krein [11, ch. V]. Our scheme is essentially the same as that developed by Krein, although the initial formulations differ since he deals directly with ODE in Banach space.

The letter N will always denote a positive integer; if it appears as a superscript in the definition of a term, it will be omitted from the notation when no confusion will result. We assume further that $[a, b] = [a, a + \kappa r]$ for some positive integer κ. More will be said in Section 8 concerning this last assumption.

For a given N, partition $[a, b]$ into κN subintervals of length r/N by $\{t_i^N\}_{i=0}^{\kappa N}$, where $t_i^N = a + ir/N$. The approximation scheme is suggested by the heuristic argument:

$$y(t_i) = \zeta + \sum_{j=0}^{i-1} \int_{t_j}^{t_{j+1}} [\mathscr{A}(s)\, y(s) + (f(s), 0)]\, ds$$

implies that

$$y(t_i) - (r/N) \sum_0^{i-1} \mathscr{A}_N(j)\, y(t_j) \approx \zeta + \sum_0^{i-1} \int_{t_j}^{t_{j+1}} (f(s), 0)\, ds,$$

where $(r/N)\, \mathscr{A}_N(j)\, y(t_j)$ approximates $\int_{t_j}^{t_{j+1}} \mathscr{A}(t)\, y(t)\, dt$ for each value of j.

Some preliminary definitions are required to discuss factor convergence. For a given N, let

$$Z_N = \bigtimes_0^N R^n \qquad \text{with inner product} \quad \langle \cdot, \cdot \rangle_{Z_N} = \langle \cdot, \cdot \rangle_{R^n} + (r/N) \sum_1^N \langle \cdot, \cdot \rangle_{R^n},$$

$$E_N = \bigtimes_0^{\kappa N} Z_N \qquad \text{with maximum norm,}$$

$$J_j^N = [-jr/N, -(j-1)\,r/N) \qquad \text{for} \quad j = 1, 2, \ldots, N,$$

$$J_0^N = \{0\},$$

$$K_i^N = [a + ir/N, a + (i+1)\,r/N) \qquad \text{for} \quad i = 0, 1, \ldots, \kappa N - 1 \qquad \text{and}$$

$$K_{\kappa N}^N = \{a + \kappa r\}.$$

Given an element $\phi \in L_2(-r, 0; R^n)$, let

$$\phi_j^N = (N/r) \int_{J_j} \phi(\theta)\, d\theta.$$

Let $E_N{}^0$ denote the Banach space of functions from $[a, b]$ to Z which are uniformly continuous on each interval K_i^N, with supremum norm $|\cdot|_{E_N{}^0}$. Define the operators

$$\pi_N\colon Z \to Z_N \text{ by } \pi_N(\eta, \phi) = (\eta, \phi_1{}^N, \ldots, \phi_N{}^N), \qquad \text{and}$$

$$p_N\colon E_N{}^0 \to E_N \text{ by } p_N y = (\pi_N y(t_0), \ldots, \pi_N y(t_{\kappa N})).$$

Note that $|\pi_N \zeta|_{Z_N} \leqslant |\zeta|_Z$ and $|p_N y|_{E_N} \leqslant |y|_{E_N{}^0}$ for all $\zeta \in Z$, $y \in E_N{}^0$.

Now define the operators

$$\pi_N^{-1}\colon Z_N \to Z \qquad \text{by} \quad \pi_N^{-1} v = \left(v_0, \sum_{j=1}^N v_j \operatorname{ch}(J_j, [-r, 0]) \right), \qquad \text{and}$$

$$p_N^{-1}\colon E_N \to E_N{}^0 \qquad \text{by} \quad p_N^{-1} w = \sum_{i=0}^{\kappa N} (\pi_N^{-1} w_i) \operatorname{ch}(K_i, [a, b]).$$

Observe that π_N^{-1}, p_N^{-1} are right inverses of π_N, p_N respectively, and that $E \subset E_N{}^0$ for all N. Note also that in general, $(p_N^{-1} w)(\cdot)$ is not continuous on $[a, b]$, so $y \in E$ does not imply that $p_N^{-1} p_N y \in R$.

The following two lemmas show us the sense in which the spaces Z_N, E_N approximate Z, E respectively.

4.1. Lemma. *For all* $\zeta \in Z$,

$$\pi_N^{-1} \pi_N \zeta \to \zeta \qquad \textit{in } Z \textit{ as} \quad N \to \infty.$$

Proof (see [3, Lemma 3.2]). This is easy to verify by computation if $\zeta = (\eta, \phi)$ and ϕ is continuous. Then use the density of the continuous functions plus the linearity and uniform equicontinuity of the maps $\zeta \to \pi_N^{-1}\pi_N\zeta$. ∎

4.2. Lemma. *For all* $y \in E$

$$\lim_{N\to\infty} | y - p_N^{-1}p_N y |_{E_N^0} = 0.$$

Proof. For $i = 0, 1,..., \kappa N$ we have $(p_N^{-1}p_N y)(t) = \pi_N^{-1}\pi_N y(t)$ when $t \in K_i$. Observe that the operators $\pi_N^{-1}\pi_N$ converge strongly to the identity in Z (Lemma 4.1); therefore $\pi_N^{-1}\pi_N\xi \to \xi$ uniformly with respect to ξ in the compact set $\{y(t): t \in [a, b]\} \subset Z$. This fact and the uniform continuity of y imply that $(p_N^{-1}p_N y)(t)$ converges to $y(t)$ uniformly with respect to $t \in [a, b]$. ∎

If $\{y_N\}_{N=1}^{\infty}$ is a sequence with the property that $y_N \in E_N$ for all N, we say that $\{y_N\}$ *factor converges* to $y \in E$ if

$$\lim_{N\to\infty} | p_N y - y_N |_{E_N} = 0.$$

Such limits, if they exist, are unique by Lemma 4.2.

Let $\gamma = (\zeta, \lambda, f)$. Recall that by Theorem 3.15, $z(t, \gamma) = (x(t, a, \gamma), x_t(a, \gamma))$ is the unique solution of Eq. (4.1). Suppose that there is a sequence $\{z_N(\gamma)\}$ which factor converges to z. For each N, define $x_N(t) = x_N(t, \gamma)$ as the first component of $p_N^{-1}z_N(\gamma)(t)$. The significance of factor convergence lies in the fact that Lemma 4.2 implies

$$\lim_{N\to\infty} \sup\{| x(t) - x_N(t)|: t \in [a, b]\} = 0.$$

Therefore, the approximation of a solution of Eq. (4.1) in fact also yields an approximation of a solution of Eqs. (2.1), (2.2).

5. The Approximation Scheme

Equation (4.1) may be rewritten in a more convenient form. To this end, let

$$E_W = \{z \in E: z(t) \in W, \forall t \text{ and } \mathscr{A}z \in E\},$$

$T = T(\lambda): E_W \to E$ be given by

$$(Ty)(t) = y(t) - \int_a^t \mathscr{A}(s)\, y(s)\, ds, \qquad \text{and}$$

$S: W \times F \to E$ be given by

$$S(\zeta, f)(t) = \zeta + \int_a^t (f(s), 0)\, ds.$$

We want to approximate solutions of

$$Ty = S(\zeta, f). \tag{5.1}$$

Therefore consider the equations

$$T_N y_N = S_N q_N(\zeta, f) \tag{5.2}$$

where

$F_N = \bigtimes_1^{\kappa N} Z_N$ with inner product

$$\langle \cdot, \cdot \rangle_{F_N} = (r/N) \sum_1^{\kappa N} \langle \cdot, \cdot \rangle_{Z_N},$$

$q_N \colon Z \times F \to Z_N \times F_N$ is given by

$$q_N(\zeta, f) = \left(\pi_N \zeta, (N/r) \int_{K_0} (f(s), 0)\, ds, \ldots, (N/r) \int_{K_{\kappa N - 1}} (f(s), 0)\, ds\right),$$

$S_N \colon Z_N \times F_N \to F_N$ is given by ($w = (w_0, \ldots, w_{\kappa N})$)

$$(S_N w)_i = \begin{cases} w_0 & i = 0 \\ w_0 + (r/N) \sum_{j=1}^{i} w_j & i = 1, 2, \ldots, \kappa N, \end{cases}$$

and $T_N \colon E_N \to E_N$ approximates T in the sense to be described below. Observe that $p_N S(\zeta, f) = S_N q_N(\zeta, f)$ for all $(\zeta, f) \in W \times F$ and that S_N^{-1} exists for all N.

The properties of the operators T_N determine the degree of success of the resulting approximation scheme. To illustrate, it is quite natural to ask that T_N^{-1} exist for all N, because this would imply existence and uniqueness of solutions of Eq. (5.2). We could ask further that the condition $p_N T y = T_N p_N y$ obtain for all y, N. This would imply that for solutions y, y_N of Eqs. (5.1), (5.2),

$$\begin{aligned} p_N y &= T_N^{-1} T_N p_N y = T_N^{-1} p_N T y = T_N^{-1} p_N S(\zeta, f) = T_N^{-1} S_N q_N(\zeta, f) \\ &= y_N \end{aligned}$$

for each N, i.e. immediate factor convergence. The first requirement is attainable, but we do not know of a scheme which satisfies the second requirement for general $\lambda \in \Lambda_c$.

The particular definitions now given lead to a scheme which has previously been studied; more will be said in Section 14 concerning this. For a given N, let

$\ell(\cdot, N) \colon \{0, 1, \ldots, \nu\} \to \{0, 1, \ldots, N\}$ be such that

$$-\tau_j \in J_{\ell(j, N)} \qquad \text{for all } j,$$

$A(N, \cdot) = A(N, \cdot, \lambda)\colon \{0, 1,..., \kappa N - 1\} \times \{0, 1,..., \nu\} \to R^{n\times n}$ be given by

$$A(N, i, j) = (N/r) \int_{K_i} A_j(t)\, dt,$$

$D(N, \cdot) = D(N, \cdot, \lambda)\colon \{0, 1,..., \kappa N - 1\} \times \{1, 2,..., N\} \to R^{n\times n}$ be given by

$$D(N, i, j) = (N/r)^2 \int_{K_i} \int_{J_j} D(t, \theta)\, d\theta\, dt,$$

$L(N, \cdot) = L(N, \cdot, \lambda)\colon \{0, 1,..., \kappa N - 1\} \to \mathscr{B}(Z_N, R^n)$ be given by

$$L(N, i)\, v = \sum_{j=0}^{\nu} A(N, i, j)\, v_{\ell(j,N)} + (r/N) \sum_{j=1}^{N} D(N, i, j)\, v_j$$

$\mathscr{A}_N = \mathscr{A}_N(\lambda)\colon \{0, 1,..., \kappa N - 1\} \to \mathscr{B}(Z_N)$ be given by

$$(\mathscr{A}_N(i)\, v)_j = \begin{cases} L(N, i)\, v & j = 0 \\ (N/r)(v_{j-1} - v_j) & j = 1, 2,..., N, \end{cases}$$

and finally let

$$T_N = T_N(\lambda)\colon E_N \to E_N \text{ be given by}$$

$$(T_N y)_i = \begin{cases} y_0 & i = 0 \\ y_i - (r/N) \sum_{j=0}^{i-1} \mathscr{A}_N(j)\, y_j & i = 1, 2,..., \kappa N. \end{cases}$$

Assuming for the moment that T_N^{-1} exists, define $z_N = T_N^{-1} S_N q_N(\zeta, f)$. Then

$$\begin{aligned} p_N z - z_N &= T_N^{-1} S_N S_N^{-1} T_N p_N z - T_N^{-1} S_N q_N(\zeta, f) \\ &= (T_N^{-1} S_N)[S_N^{-1} T_N p_N z - q_N(\zeta, f)]. \end{aligned}$$

Hence factor convergence is established once we demonstrate that:

(i) $\sup\{| T_N^{-1} S_N |_{\mathscr{B}(Z_N \times F_N, E_N)}\colon N = 1, 2,...\} < \infty$ (this is known as *stability* of the difference scheme) and

(ii) $\lim_N | S_N^{-1} T_N p_N z - q_N(\zeta, f)|_{Z_N \times F_N} = 0$ (this is known as *consistency* of the difference scheme).

Stability and consistency will be proved in Sections 6 and 7 respectively.

It might appear more natural to write $p_N z - z_N = (T_N^{-1})[T_N p_N z - S_N q_N(\zeta, f)]$ and establish the corresponding stability and consistency results. However, we have been unable to demonstrate that in this case the stability condition obtains for general systems.

The following lemma is given for reference; its proof is straightforward (use induction for part (i)). Note the similarity of part (iii) with the function given in Krein [11, p. 342].

5.1. LEMMA. *With S_N, T_N as above, we have:*

(i) T_N^{-1}: $E_N \to E_N$ *given by*

$$(T_n^{-1}v)_i = \begin{cases} v_0 & i = 0, \\ [I + (r/N)\mathscr{A}_N(0)]\, v_0 + (v_1 - v_0) & i = 1, \\ \prod_{j=0}^{i-1} [I + (r/N)\mathscr{A}_N(j)]\, v_0 + \sum_{k=0}^{i-2} \left\{ \prod_{j=k+1}^{i-1} [I + (r/N)\mathscr{A}_N(j)] \right\} (v_{k+1} - v_k) & \\ \qquad + (v_i - v_{i-1}) & i = 2, 3, \ldots, kN; \end{cases}$$

(ii) S_N^{-1}: $E_N \to Z_N \times F_N$ *given by*

$$(S_N^{-1}v)_i = \begin{cases} v_0 & i = 0 \\ (N/r)(v_i - v_{i-1}) & i = 1, 2, \ldots, \kappa N; \end{cases}$$

(iii) $T_N^{-1}S_N$: $Z_N \times F_N \to E_N$ *given by*

$$(T_N^{-1}S_N v)_i = \begin{cases} v_0 & i = 0 \\ [I + (r/N)\mathscr{A}_N(0)]\, v_0 + (r/N)\, v_1 & i = 1 \\ \prod_{j=0}^{i-1} [I + (r/N)\mathscr{A}_N(j)]\, v_0 + (r/N) \sum_{k=0}^{i-2} \left\{ \prod_{j=k+1}^{i-1} [I + (r/N)\mathscr{A}_N(j)] \right\} v_{k+1} & \\ \qquad + (r/N)\, v_i & i = 2, 3, \ldots, \kappa N; \end{cases}$$

(iv) $S_N^{-1}T_N$: $E_N \to Z_N \times F_N$ *given by*

$$(S_N^{-1}T_N v)_i = \begin{cases} v_0 & i = 0, \\ (N/r)(v_i - [I + (r/N)\mathscr{A}_N(i-1)]\, v_{i-1}) & i = 1, 2, \ldots, \kappa N. \end{cases}$$

6. STABILITY

In proving that the desired stability condition indeed obtains, we find as in the proof of Lemma 3.3 that it is more convenient to work in certain spaces with equivalent inner product topologies. So for a given N, let

$k(\cdot, N)$: $\{0, 1, \ldots, N-1\} \to \{1, 2, \ldots, \nu\}$ be given by

$$k(j, N) = \min \left\{ k \geqslant 1 : -\tau_k \in \bigcup_{i=j+1}^{N} J_i \right\},$$

$$\beta_j = \beta_j^N(\lambda) = \begin{cases} 1 & j = N \\ 1 + \sum_{i=k(j,N)}^{\nu} |A_i|_c & j = 0, 1, 2, \ldots, N-1, \end{cases}$$

and

$$Y_N = Y_N(\lambda) = \bigtimes_0^N R^n \text{ with inner product}$$

$$\langle \cdot, \cdot \rangle_{Y_N} = \langle \cdot, \cdot \rangle_{R^n} + (r/N) \sum_1^N \beta_{j-1} \langle \cdot, \cdot \rangle_{R^n} .$$

For $\lambda \in A_c$, define $M(\lambda) = (1 + \sum_1^{\nu} | A_j |_c)^{1/2}$. Observe that $1 \leqslant \beta_0^N \leqslant M^2(\lambda)$ for all N. Thus the new norms satisfy

$$| \cdot |_{Z_N} \leqslant | \cdot |_{Y_N} \leqslant M(\lambda) | \cdot |_{Z_N} .$$

One purpose for these rather complicated definitions is that they enable us in the proof of Lemma 6.2 below to write for $v \in Z_N$,

$$\sum_{j=1}^{\nu} | A_j |_c | v_{\ell(j,N)} | = \sum_{j=1}^{N} (\beta_{j-1} - \beta_j) | v_j |.$$

Define the spaces

$$\mathscr{E}_N = \mathscr{E}_N(\lambda) = \bigtimes_0^{\kappa N} Y_N \qquad \text{with maximum norm,}$$

$$\mathscr{F}_N = \mathscr{F}_N(\lambda) = \bigtimes_1^{\kappa N} Y_N \qquad \text{with inner product}$$

$$\langle \cdot, \cdot \rangle_{\mathscr{F}_N} = (r/N) \sum_1^{\kappa N} \langle \cdot, \cdot \rangle_{Y_N} .$$

Then

$$| \cdot |_{E_N} \leqslant | \cdot |_{\mathscr{E}_N} \leqslant M(\lambda) | \cdot |_{E_N} , \qquad \text{and}$$
$$| \cdot |_{F_N} \leqslant | \cdot |_{\mathscr{F}_N} \leqslant M(\lambda) | \cdot |_{F_N} .$$

We will show that

$$\sup\{| T_N^{-1} S_N |_{\mathscr{B}(Y_N \times \mathscr{F}_N, \mathscr{E}_N)} : N = 1, 2, \ldots\} < \infty,$$

which will yield the desired stability result.

Lemmas 6.1 and 6.2 contain the crux of the stability argument, which is summarized in Theorem 6.3. The idea of defining the inner product used in the proof of Lemma 6.2 was motivated by a similar definition in [16]. Define the symbol ρ as $\rho = \max\{1, r^{1/2}\}$.

6.1. LEMMA. *Suppose that for a given* $\lambda \in \Lambda_c$ *there is an* $\alpha = \alpha(\lambda) \geqslant 0$ *for which* $\max\{| I + (r/N)\mathscr{A}_N(i)|_{\mathscr{B}(Y_N)}: 0 \leqslant i \leqslant \kappa N - 1\} \leqslant 1 + \alpha r/N$. *Then there is a constant* $\beta = \beta(\lambda)$ *such that for all* N,

$$| T_N^{-1} S_N |_{\mathscr{B}(Y_N \times \mathscr{F}_N, \mathscr{E}_N)} \leqslant \beta e^{\alpha\kappa r}.$$

Proof. The existence of such an α implies that for all N and all j, k with $0 \leqslant k \leqslant j \leqslant \kappa N - 1$,

$$\left| \prod_{i=k}^{j} [I + (r/N)\mathscr{A}_N(i)] \right|_{\mathscr{B}(Y_N)} \leqslant e^{\alpha\kappa r}.$$

Consequently for all $i = 0, 1,..., \kappa N$ (see Lemma 5.1(iii))

$$\begin{aligned} |(T_N^{-1} S_N v)_i |_{Y_N} &\leqslant e^{\alpha\kappa r} \left(| v_0 |_{Y_N} + (r/N) \sum_{j=1}^{i} | v_j |_{Y_N} \right) \\ &\leqslant 2e^{\alpha\kappa r} \left(| v_0 |_{Y_N}^2 + (\kappa r)(r/N) \sum_{j=1}^{\kappa N} | v_j |_{Y_N}^2 \right)^{1/2} \\ &\leqslant 2\kappa^{1/2} \rho e^{\alpha\kappa r} | v |_{Y_N \times \mathscr{F}_N}. \quad \blacksquare \end{aligned}$$

6.2. LEMMA. *Given* $\lambda \in \Lambda_c$ *there is an* $\alpha = \alpha(\lambda)$ *for which*

$$\max\{| I + (r/N)\mathscr{A}_N(i)|_{\mathscr{B}(Y_N)}: 0 \leqslant i \leqslant \kappa N - 1\} \leqslant 1 + \alpha r/N.$$

Proof. For $v \in Y_N$,

$$([I + (r/N)\mathscr{A}_N(i)]\, v)_k = \begin{cases} v_0 + (r/N) L(N, i)\, v & k = 0 \\ v_{k-1} & k = 1, 2,..., N. \end{cases}$$

Hence,

$$\begin{aligned} &|([I + (r/N)\mathscr{A}_N(i)]\, v)_0 |_{R^n} \\ &\leqslant | v_0 | + (r/N) \sum_{j=0}^{\nu} | A_j |_c \, | v_{\ell(j,N)} | + (r/N)^2 \sum_{j=1}^{N} | D(N, i, j)| \, | v_j | \\ &\leqslant | v_0 | + (r/N) \, | A_0 |_c \, | v_0 | + (r/N) \sum_1^N (\beta_{j-1} - \beta_j) \, | v_j | + (r/N) \, | D |_c \, | v |_{Z_N} \\ &\leqslant | v_0 | + (r/N) \left\{ | \lambda |_c \, | v |_{Y_N} + \sum_1^N (\beta_{j-1} - \beta_j) \, | v_j | \right\}. \end{aligned}$$

In the above series of inequalities we have used

$$(r/N)\sum_{j=1}^{N} |\, D(N,i,j)|\,|\, v_j\,| \leqslant \left[(r/N)\sum_{j=1}^{N} |\, D(N,i,j)|^2\right]^{1/2} |\, v\,|_{Z_N}$$

where

$$(r/N)\sum_{j=1}^{N} |\, D(N,i,j)|^2 = (r/N)\sum_{j=1}^{N} \left|(N/r)^2 \int_{K_i}\int_{J_j} D(t,\theta)\,d\theta\,dt\right|^2$$

$$\leqslant (N/r)\int_{K_i}\sum_{j=1}^{N}\int_{J_j} |\, D(t,\theta)|^2\,d\theta\,dt$$

$$\leqslant |\, D\,|_c^2\,.$$

Thus letting $\{\cdots\} = \{|\,\lambda\,|_c\,|\,v\,|_{Y_N} + \sum_1^N (\beta_{j-1} - \beta_j)\,|\,v_j\,|\}$, we have

$$|([I + (r/N)\mathscr{A}_N(i)]\,v)_0\,|_{R^n}^2 \leqslant |\,v_0\,|^2 + 2(r/N)\,|\,v_0\,|\,\{\cdots\} + (r/N)^2\{\cdots\}^2.$$

Using the inequality $2cd \leqslant c^2 + d^2$, we find that the middle term on the right-hand side satisfies

$$\begin{aligned} 2(r/N)\,|\,v_0\,|\{\cdots\} &= 2(r/N)[(2\beta_0)^{1/2}\,|\,v_0\,|][(1/2\beta_0)^{1/2}\{\cdots\}] \\ &\leqslant 2\beta_0(r/N)\,|\,v_0\,|^2 + (1/2\beta_0)(r/N)\{\cdots\}^2. \end{aligned}$$

Observe that

$$\{\cdots\}^2 \leqslant 2\,|\,\lambda\,|_c^2\,|\,v\,|_{Y_N}^2 + 2\left\{\sum_1^N (\beta_{j-1} - \beta_j)\,|\,v_j\,|\right\}^2$$

$$\leqslant 2\,|\,\lambda\,|_c^2\,|\,v\,|_{Y_N}^2 + 2\left\{\sum_1^N (\beta_{j-1} - \beta_j)\right\}\sum_1^N (\beta_{j-1} - \beta_j)\,|\,v_j\,|^2$$

$$\leqslant 2\,|\,\lambda\,|_c^2\,|\,v\,|_{Y_N}^2 + 2\beta_0\sum_1^N (\beta_{j-1} - \beta_j)\,|\,v_j\,|^2.$$

Recalling that $\beta_0 \geqslant 1$,

$$2(r/N)\,|\,v_0\,|\,\{\cdots\}$$

$$\leqslant 2\beta_0(r/N)\,|\,v_0\,|^2 + |\,\lambda\,|_c^2\,(r/N)\,|\,v\,|_{Y_N}^2 + (r/N)\sum_1^N (\beta_{j-1} - \beta_j)\,|\,v_j\,|^2$$

$$\leqslant (r/N)[2\beta_0 + |\,\lambda\,|_c^2]\,|\,v\,|_{Y_N}^2 + (r/N)\sum_1^N (\beta_{j-1} - \beta_j)\,|\,v_j\,|^2$$

and

$$(r/N)^2 \{\cdots\}^2 \leqslant (r/N)[2r \mid \lambda \mid_c^2 \mid v \mid_{Y_N}^2 + 2\beta_0 \mid v \mid_{Y_N}^2].$$

Consequently,

$$\begin{aligned}
&|[I + (r/N) \mathscr{A}_N(i)] v |_{Y_N}^2 \\
&\quad \leqslant | v_0 |^2 + (r/N)[(1 + 2r) \mid \lambda \mid_c^2 + 4\beta_0] \mid v \mid_{Y_N}^2 \\
&\qquad + (r/N) \sum_1^N (\beta_{j-1} - \beta_j) \mid v_j \mid^2 + (r/N) \sum_1^N \beta_{j-1} \mid v_{j-1} \mid^2 \\
&\quad \leqslant | v_0 |^2 + (r/N)[(1 + 2r) \mid \lambda \mid_c^2 + 5\beta_0] \mid v \mid_{Y_N}^2 + (r/N) \sum_1^N \beta_{j-1} \mid v_j \mid^2 \\
&\quad \leqslant (1 + \alpha r/N) \mid v \mid_{Y_N}^2 ,
\end{aligned}$$

where $\alpha = (1 + 2r) \mid \lambda \mid_c^2 + 5M^2(\lambda)$. Therefore

$$|[I + (r/N) \mathscr{A}_N(i)] v |_{Y_N} \leqslant (1 + \alpha r/N)^{1/2} \mid v \mid_{Y_N} \leqslant (1 + \alpha r/N) \mid v \mid_{Y_N} . \quad \blacksquare$$

6.3. Theorem. *Suppose* $G_1 \subset \Lambda_c$ *is such that* $\sup\{| \lambda |_c : \lambda \in G_1\} < \infty$. *Then*

$$\sup\{| T_N^{-1}(\lambda) S_N |_{\mathscr{B}(Z_N \times F_N, E_N)} : \lambda \in G_1 ; N = 1, 2,\ldots\} < \infty.$$

Proof. Note that $\sup\{M(\lambda) : \lambda \in G_1\} < \infty$ and

$$\begin{aligned}
&\sup\{| T_N^{-1}(\lambda) S_N |_{\mathscr{B}(Y_N \times \mathscr{F}_N(\lambda), \mathscr{E}_N(\lambda))} : \lambda \in G_1 ; N = 1, 2,\ldots\} \\
&\qquad \leqslant \sup\{\beta(\lambda) e^{\alpha(\lambda)\kappa r} : \lambda \in G_1\} < \infty.
\end{aligned}$$

The conclusion follows immediately. $\blacksquare$

7. Consistency

As usual, for $\gamma = (\zeta, \lambda, f)$ let $x = x(a, \gamma)$ denote the solution of Eqs. (2.1), (2.2); let $z_N = z_N(\gamma) = T_N^{-1}(\lambda) S_N q_N(\zeta, f)$. We shall restrict our attention to $\gamma \in G$, where $G = \{\zeta\} \times G_1 \times G_2$ is such that: $\zeta \in W$; $G_1 \subset \Lambda_c$ is relatively compact in Λ with $m_1 = \sup\{| \lambda |_c : \lambda \in G_1\}$ finite; and $G_2 \subset F$ is bounded. It will be of interest later to know that $\{z_N(\gamma)\}$ factor converges to $z(\gamma)$ uniformly with respect to $\gamma \in G$ (the function z was defined in Section 2). Thus the primary result of this section is the statement of Theorem 7.5 that the consistency condition is fulfilled uniformly with respect to $\gamma \in G$.

By definition of the operators p_N and q_N,

$$[S_N^{-1}T_N p_N z(\gamma)]_i$$
$$= \begin{cases} \pi_N\zeta & i = 0 \\ (N/r)\{\pi_N z(t_i, \gamma) - [I + (r/N)\mathscr{A}_N(i-1)]\,\pi_N z(t_{i-1}, \gamma)\} & i = 1, 2,..., \kappa N, \end{cases}$$

and

$$[q_N(\zeta, f)]_i = \begin{cases} \pi_N\zeta & i = 0 \\ (N/r)\int_{K_{i-1}} (f(s), 0)\, ds & i = 1, 2,..., \kappa N. \end{cases}$$

Hence, defining

$$g(N, \gamma)(i) = (r/N)[S_N^{-1}T_N p_N z(\gamma) - q_N(\zeta, f)]_i,$$

we obtain

$$g(N, \gamma)(i$$
$$= \begin{cases} 0 \quad i = 0 \\ \pi_N z(t_i, \gamma) - \left\{[I + (r/N)\mathscr{A}_N(i-1)]\,\pi_N z(t_{i-1}, \gamma) + \int_{K_{i-1}} (f(s), 0)\, ds\right\} \\ \qquad i = 1, 2,..., \kappa N. \end{cases}$$

Observe that $g(N, \gamma)(i) \in Z_N$; let $g(N, \gamma)(i, j)$ denote its jth component.

Taking advantage of the representation $z(t, \gamma) = (x(t, a, \gamma), x_t(a, \gamma))$, we find that for $j = 2, 3,..., N$

$$\begin{aligned} g(N, \gamma)(i, j) &= (\pi_N z(t_i, \gamma))_j - (\pi_N z(t_{i-1}, \gamma))_{j-1} \\ &= (x_{t_i}(a, \gamma))_j^N - (x_{t_{i-1}}(a, \gamma))_{j-1}^N \\ &= 0. \end{aligned}$$

The terms $g(N, \gamma)(i, j)$ for $j = 0, 1$ must be analyzed separately and in more detail.

7.1. Lemma. (i) $m_2 = \sup\{|\, x_t(a, \gamma)|_c\colon (t, \gamma) \in [a, b] \times G\}$ *is finite*;

(ii) $\Gamma = \sup\{|\, d/dt\, x(\cdot, a, \gamma)|_{L_2(a-r,b)}\colon \gamma \in G\}$ *is finite.*

Proof. Let $(\eta, \phi) = \zeta$. Observe that $|\, x_a(a, \gamma)|_c = |\,\phi\,|_c$ for all γ and that $\sup\{|\, x(t; a, \gamma)|\colon a \leqslant t \leqslant b, \gamma \in G\}$ is finite by the hypotheses on G_1, G_2 and Lemma 2.2, Theorems 3.5, 3.14. Consequently $x(\gamma)$ is continuous on $[a - r, b]$

and bounded uniformly with respect to γ, so m_2 is finite. Condition (ii) follows from the relations (recall that $\rho = \max\{1, r^{1/2}\}$):

$$\begin{aligned} |\dot{x}(t)| &\leqslant \rho \mid \lambda \mid_c \mid x_t \mid_c + |f(t)| \\ &\leqslant \rho m_1 m_2 + |f(t)| \end{aligned}$$

for $t \in [a, b]$ and the fact that for all γ, $|\dot{x}_a(a, \gamma)|_{L_2} = |\dot{\phi}|_{L_2}$. ∎

Define

$$h(i, \gamma) = \left\{\int_{K_{i-1}} |\dot{x}(t, a, \gamma)|^2 \, dt\right\}^{1/2}$$

for $i = 1, 2,\ldots, \kappa N$. The following lemma is easily proved using Hölder's inequality and the definitions of $g(N, \gamma)(i, 1)$ and $h(i, \gamma)$.

7.2. Lemma. $\sup\{|g(N, \gamma)(i, 1)|: 1 \leqslant i \leqslant \kappa N, \gamma \in G\} \leqslant h(i, \gamma)(r/N)^{1/2}$.

The analysis of $g(N, \gamma)(i, 0)$ utilizes the following elementary result.

7.3. Lemma. *For all $t, s \in [a - r, b]$ and all $\gamma \in G$ we have $|x(t) - x(s)| \leqslant \Gamma |t - s|^{1/2}$. Furthermore, given $i \in \{1, 2,\ldots, \kappa N\}$, $j \in \{1, 2,\ldots, N\}$ and $t \in [a + (i - 1 - j)r/N, a + (i + 1 - j)r/N]$, we have*

$$|x(t, a, \gamma) - (x_{t_{i-1}}(a, \gamma))_j^N| \leqslant 2\Gamma(r/N)^{1/2}.$$

Proof. The first statement is an immediate consequence of Lemma 7.1. Let $\tau = a + (i - 1 - j)r/N$ and $\sigma = t - \tau$. Then $\sigma \in [0, 2r/N]$ by hypothesis. Observe that

$$\begin{aligned} |x(t) - (x_{t_{i-1}})_j^N| &= \left| x(\tau + \sigma) - (N/r) \int_\tau^{\tau+(r/N)} x(\theta)\, d\theta \right| \\ &= (N/r) \left| \int_0^{r/N} [x(\tau + \sigma) - x(\tau + \theta)]\, d\theta \right| \\ &\leqslant (N/r) \int_0^{r/N} \Gamma \mid \sigma - \theta \mid^{1/2} d\theta. \end{aligned}$$

The function $\sigma \to \int_0^{r/N} |\sigma - \theta|^{1/2} \, d\theta$ assumes the maximum value of $(2/3)(2^{3/2} - 1)(r/N)^{3/2} \leqslant 2(r/N)^{3/2}$ (when $\sigma = 2r/N$) over the interval $[0, 2r/N]$. ∎

7.4. Lemma. *For all $\gamma \in G$, $\max\{|g(N, \gamma)(i, 0)|: 1 \leqslant i \leqslant \kappa N\} \leqslant 2\rho\Gamma m_1(r/N)^{3/2}$.*

Proof. Observe that

$$\begin{aligned} g(N,\gamma)(i,0) &\\ &= x(t_i) - \left\{x(t_{i-1}) + (r/N)L(N,i-1)\,\pi_N z(t_{i-1}) + \int_{K_{i-1}} f(t)\,dt\right\} \\ &= \int_{K_{i-1}} [L(t,x_t) + f(t)]\,dt - (r/N)L(N,i-1)\,\pi_N z(t_{i-1}) - \int_{K_{i-1}} f(t)\,dt \\ &= \int_{K_{i-1}} L(t,x_t)\,dt - (r/N)L(N,i-1)\,\pi_N z(t_{i-1}). \end{aligned}$$

Thus

$$\begin{aligned} &| g(N,\gamma)(i,0)| \\ &= \Big| \sum_{j=0}^{\nu} \int_{K_{i-1}} A_j(t)\,x(t-\tau_j) + \int_{K_{i-1}} \sum_{j=1}^{N} \int_{J_j} D(t,\theta)\,x(t+\theta)\,d\theta\,dt \\ &\quad - \int_{K_{i-1}} A_0(t)\,x(t_{i-1})\,dt - \sum_{j=1}^{\nu} \int_{K_{i-1}} A_j(t)(x_{t_{i-1}})^N_{\ell(j,N)}\,dt \\ &\quad - \int_{K_{i-1}} \sum_{j=1}^{N} \int_{J_j} D(t,\theta)(x_{t_{i-1}})^N_j\,d\theta\,dt \Big| \\ &\leqslant \int_{K_{i-1}} | A_0 |_c\, | x(t) - x(t_{i-1})|\,dt + \sum_{j=1}^{\nu} \int_{K_{i-1}} | A_j |_c\, | x(t-\tau_j) - (x_{t_{i-1}})^N_{\ell(j,N)} |\,dt \\ &\quad + \int_{K_{i-1}} \sum_{j=1}^{N} \int_{J_j} | D(t,\theta)|\, | x(t+\theta) - (x_{t_{i-1}})^N_j |\,d\theta\,dt \\ &\leqslant 2\Gamma(r/N)^{1/2} \left\{ \sum_{j=0}^{\nu} \int_{K_{i-1}} | A_j |_c\,dt + \int_{K_{i-1}} \int_{-r}^{0} | D(t,\theta)|\,d\theta\,dt \right\} \\ &\leqslant 2\Gamma(r/N)^{1/2} \left\{ (r/N) \sum_{j=0}^{\nu} | A_j |_c + r^{1/2} \int_{K_{i-1}} | D(t)|_{L_2}\,dt \right\} \\ &\leqslant 2\rho\Gamma m_1 (r/N)^{3/2}. \quad \blacksquare \end{aligned}$$

The consistency of the approximation scheme is now readily established.

7.5. THEOREM. *The limit*

$$0 = \lim_N | S_N^{-1} T_N p_N z(\gamma) - q_N(\zeta, f)|_{Z_N \times F_N}$$

exists uniformly with respect to $\gamma \in G$.

Proof. By Lemmas 7.1, 7.2, 7.4 and preceding remarks,

$$\begin{aligned}
&| S_N^{-1} T_N p_N z(\gamma) - q_N(\zeta, f)|^2_{Z_N \times F_N} \\
&\quad = (r/N) \sum_{i=1}^{\kappa N} (N/r)^2 \,|\, g(N, \gamma)(i)|^2_{Z_N} \\
&\quad = (N/r) \sum_{i=1}^{\kappa N} \{| g(N, \gamma)(i, 0)|^2 + (r/N) \,|\, g(N, \gamma)(i, 1)|^2\} \\
&\quad \leqslant (N/r) \sum_{i=1}^{\kappa N} \{(2\rho \Gamma m_1)^2 \,(r/N)^3 + h^2(i, \gamma)(r/N)^2\} \\
&\quad = \kappa N (r/N)^2 \,(2\rho\Gamma m_1)^2 + (r/N) \sum_{i=1}^{\kappa N} \int_{K_{i-1}} | \dot{x}(t)|^2 \, dt \\
&\quad \leqslant (r/N)\, \Gamma^2 [1 + \kappa r (2\rho m_1)^2].
\end{aligned}$$

Therefore

$$| S_N^{-1} T_N p_N z(\gamma) - q_N(\zeta, f)|_{Z_N \times F_N} \leqslant m_3 (r/N)^{1/2},$$

where $m_3 = \Gamma[1 + \kappa r(2\rho m_1)^2]^{1/2}$ is independent of $\gamma \in G$. ∎

8. Approximation Under More General Assumptions

So far, we have relied greatly on the fact that λ is continuously differentiable. As we shall see below, the preceding results enable us to approximate solutions of Eqs. (2.1), (2.2) whenever $\lambda \in \Lambda$ is essentially bounded. Once this has been established, we may deal with such FDE on any bounded interval $[a, b]$ by simply extending λ and f as zero to the interval $[a, a + \kappa r]$ for an appropriate integer κ.

Let

$$\Lambda_\infty = L_\infty \left(a, b; \left(\bigtimes_0^\nu R^{n \times n}\right) \times L_2(a, b; R^{n \times n})\right)$$

with norm $| \cdot |_\infty$. Using a standard result [12, Corollary, p. 288] one can prove that for every $\lambda \in \Lambda_\infty$ there is a sequence $\{\psi_i\}$ in Λ_∞ of step (i.e. finitely-valued) functions which converges to λ in the Λ norm, and is such that $| \psi_i |_\infty \leqslant 2 \,|\, \lambda |_\infty$ for all i. From $\{\psi_i\}$ one can construct a sequence $\{\lambda_i\}$ in Λ_c having the same properties, i.e. $| \lambda_i - \lambda | \to 0$ and for all i, $| \lambda_i |_c \leqslant 2 \,|\, \lambda |_\infty$.

Suppose $(\zeta, \lambda, f) \in W \times \Lambda_\infty \times F$ is given; let $\{\lambda_i\}$ be as above and define γ, γ_i $(i = 1, 2, \ldots)$ as (ζ, λ, f), (ζ, λ_i, f) respectively. Then the set $G = \bigcup_i \{\gamma_i\}$

satisfies the standing assumption of Section 7, hence also the hypotheses of Theorems 6.3 and 7.5. Therefore

$$\lim_{N} | p_N z(\gamma_N) - z_N(\gamma_N)|_{E_N} = 0.$$

Note that $z(\gamma_N)$ converges to $z(\gamma)$ in E by Lemma 2.2. Consequently $\{z_N(\gamma_N)\}$ factor converges to $z(\gamma)$, since

$$| p_N z(\gamma) - z_N(\gamma_N)|_{E_N} \leqslant | z(\gamma) - z(\gamma_N)|_E + | p_N z(\gamma_N) - z_N(\gamma_N)|_{E_N} . \quad (8.1)$$

Assume now that $f, f_i \in F$ $(i = 1, 2,...)$ and $\{f_i\} \rightharpoonup f$; redefine γ_i as (ζ, λ_i, f_i). Then Lemma 2.2 and Theorem 3.14 in conjunction with the above reasoning and standard inequalities, imply that once again $\{z_N(\gamma_N)\}$ factor converges to $z(\gamma)$. This fact will be used in the proof of Lemma 10.1.

Since we are primarily concerned with the computational aspects of this theory, it is of interest to know how the operators $\mathscr{A}_N(\lambda_N)$ relate to $\mathscr{A}_N(\lambda)$. To this end, let ϵ_0 be chosen so that $2\epsilon_0$ is the smallest positive number representable in the language to be used on a given computer (e.g. $2\epsilon_0 = 16^{-65}$ in double precision FORTRAN on IBM 360, 370 machines). If in addition to the above requirements on λ_N we ask that $| \lambda - \lambda_N | \leqslant \epsilon_0 \min\{(r/N)^{1/2}, r/N\}$, then for all j, k, N:

$$| A(N, j, k, \lambda) - A(N, j, k, \lambda_N)| \leqslant (N/r)^{1/2} | \lambda - \lambda_N | < \epsilon_0 , \qquad \text{and}$$
$$| D(N, j, k, \lambda) - D(N, j, k, \lambda_N)| \leqslant (N/r) | \lambda - \lambda_N | < \epsilon_0 .$$

Thus the machine representations of $\mathscr{A}_N(\lambda_N)$ and $\mathscr{A}_N(\lambda)$ are identical.

Observe from Lemmas 7.1 and 7.4 that the rate at which $\{p_N z(\gamma_N) - z_N(\gamma_N)\}$ factor converges to zero is determined by m_1, m_2 and the bound in F on $\{f_i\}$. The overall rate of factor convergence of $z_N(\gamma_N)$ to $z(\gamma)$ depends also (see Eq. (8.1)) on the rate of convergence in E of $z(\gamma_N)$ to $z(\gamma)$, which in turn depends on the particular sequences $\{\lambda_i\}$ and $\{f_i\}$.

9. The Control Problem

Having seen that solutions of Eqs. (2.1), (2.2) may be approximated by solutions of the difference equations (5.2), we proceed to deal with an associated optimal control problem, which will be denoted as $(\mathscr{P})$. We replace the inhomogeneous term f appearing on the right-hand side of Eq. (2.1) with the product Bu, where $B \in L_\infty(a, b; R^{n\times m})$ and $u \in L_2(a, b; R^m)$.

Throughout this chapter we shall use the letter x to denote a solution of Eqs. (2.1), (2.2). For a given N, the symbol z_N will denote a solution of Eq. (5.2); the symbol x_N will denote the corresponding first component of $p_N^{-1} z_N$ (see Section 4).

Let U denote the space of control functions, $L_2(a, b; R^m)$. Assume that functions g_1: $Z \times F \to R$ and g_2: $U \to R$ have been defined and that a subset $\mathscr{U}$ of U has been specified. Define the cost functional Φ: $\mathscr{U} \to R$ by

$$\Phi(u) = g_1(z(b, a, \zeta, \lambda, Bu), x(\cdot, a, \zeta, \lambda, Bu)) + g_2(u).$$

The optimization problem is

$$(\mathscr{P}): \quad \text{Minimize } \Phi \text{ over } \mathscr{U}.$$

Problem $(\mathscr{P})$ is tractable under the hypotheses given below. We need the concept of quasiconvexity for the statement of these hypotheses. A real-valued function h defined on a convex set is *quasiconvex* (see [4, Section 5]) if

$$h(\alpha u + (1 - \alpha)v) \leqslant \max\{h(u), h(v)\}$$

for $0 \leqslant \alpha \leqslant 1$ and all u, v. If strict inequality obtains for $0 < \alpha < 1$ and $u \neq v$, h is *strictly quasiconvex*. Equivalently (see [6, Definition 1.5.2]), h is quasiconvex if the set $\{u$: $h(u) \leqslant \alpha\}$ is convex for all real numbers α.

We assume throughout that:

(H1) $\mathscr{U}$ is closed and convex;

(H2) g_i is continuous ($i = 1, 2$);

(H3) g_i is quasiconvex ($i = 1, 2$); and

(H4) (a) $\mathscr{U}$ is bounded, or

(b) (i) g_i is bounded below ($i = 1, 2$),

(ii) g_2 is radially unbounded (i.e. $g_2 \to \infty$ as $|u| \to \infty$),

(iii) the mappings Q_N^{-1} and the sets $\mathscr{U}_N$ (defined in Section 10 below) satisfy $Q_N^{-1}\mathscr{U}_N \subset \mathscr{U}$ for all N.

Other than (H4b)(iii), the above hypotheses are standard in control theory. From Lemma 2.4 we see that the maps $u \to z(b, a, \zeta, \lambda, Bu)$ and $u \to x(\cdot, a, \zeta, \lambda, Bu)$ are affine and continuous. Consequently Φ is continuous and quasiconvex; Φ is strictly quasiconvex if g_2 is. By Mazur's theorem (see [12, p. 85]), the set $\mathscr{U}$ is weakly closed. This theorem also implies that Φ, being lower semicontinuous and quasiconvex, is weakly lower semicontinuous.

Let the sequence $\{u_i\}$ be such that $\Phi(u_i) \to \alpha = \inf\{\Phi(u)$: $u \in \mathscr{P}\}$. Hypothesis (H4) implies that $\{u_i\}$ is a bounded sequence. Hence, there is a weakly convergent subsequence $\{u_{i(j)}\}$; let u^* denote its weak limit. Clearly

$$\alpha \leqslant \Phi(u^*) \leqslant \liminf \Phi(u_{i(j)}) = \alpha.$$

Therefore, problem $(\mathscr{P})$ has a solution. The optimal control is unique if we assume that Φ is strictly quasiconvex.

As in Section 8, no problem is encountered in extending the interval on which solutions of Eq. (2.1) are to be defined. In particular, if we extend λ and u as zero over some interval $[b, a + \kappa r]$, then $z(t) = z(b)$ for $b \leqslant t \leqslant a + \kappa r$. Thus only trivial adjustments need be made in defining a new cost functional. For this reason we assume henceforth with no loss of generality that $b - a = \kappa r$ for some integer κ.

10. The Approximate Control Problems

We now associate cost functionals Φ_N with Eqs. (5.2) in such a way that the resulting optimization problems, denoted $(\mathscr{A}\mathscr{P})_N$, reflect several properties of problem $(\mathscr{P})$. To this end, define

$$U_N = \bigtimes_1^{\kappa N} R^m \qquad \text{with inner product} \quad \langle \cdot, \cdot \rangle_{U_N} = (r/N) \sum_1^{\kappa N} \langle \cdot, \cdot \rangle_{R^m} .$$

Let Q_N denote the map from U into U_N given by

$$Q_N u = \left((N/r) \int_{K_0} u(t)\, dt, \ldots, (N/r) \int_{K_{\kappa N-1}} u(t)\, dt \right).$$

Observe that $| Q_N |_{\mathscr{B}(U, U_N)} \leqslant 1$ for all N. Define a right inverse $Q_N^{-1}: U_N \to U$ of Q_N by

$$Q_N^{-1} v = \sum_{i=0}^{\kappa N-1} v_{i+1} \operatorname{ch}(K_i, [a, b]).$$

(Recall that similar definitions were made in Section 4).

Define $\mathscr{U}_N = Q_N \mathscr{U}$. Let $\Phi_N: \mathscr{U}_N \to R$ be given by

$$\Phi_N(v) = g_1((p_N^{-1} z_N)(\zeta, \lambda, BQ_N^{-1} v)(b), x_N(\zeta, \lambda, BQ_N^{-1} v)(\cdot)) + g_2(Q_N^{-1} v).$$

We define the approximate optimization problems as

$$(\mathscr{A}\mathscr{P})_N: \quad \text{Minimize } \Phi_N \text{ over } \mathscr{U}_N .$$

We could have defined Φ_N as a function from $\mathscr{U}$ to R. However, this would have changed the nature of $(\mathscr{A}\mathscr{P})_N$ from that of a classical mathematical programming problem to an optimization problem over an infinite-dimensional control space.

Since the maps $v \to (p_N^{-1} z_N)(\zeta, \lambda, BQ_N^{-1} v)(b)$ and $v \to x_N(\zeta, \lambda, BQ_N^{-1} v)(\cdot)$ are affine and continuous (obvious from Lemma 5.1(iii)) for each N, each Φ_N is a continuous and quasiconvex function over the closed convex set $\mathscr{U}_N$. Using arguments similar to those employed in Section 9, we may conclude by hypothesis (H4) that for each N there is a solution $u_N^* \in \mathscr{U}_N$ of problem $(\mathscr{A}\mathscr{P})_N$. Observe that Φ_N is strictly quasiconvex if g_2 is, in which case u_N^* is unique.

Suppose for the moment that the following results have been established:

10.1. LEMMA. *If $Q_N^{-1}u_N \rightharpoonup u$, then $\Phi(u) \leqslant \liminf \Phi_N(u_N)$.*

10.2. LEMMA. *If $Q_N^{-1}u_N \to u$, then $\Phi(u) = \lim \Phi_N(u_N)$.*

10.3. LEMMA. *For all $u \in U$, $Q_N^{-1}Q_N u \to u$.*

10.4. LEMMA. *Suppose $\{v_N\}$ is bounded in U and $v_N \rightharpoonup v$. Then $Q_N^{-1}Q_N v_N \rightharpoonup v$.*

For each N, let u_N^* denote a solution of problem $(\mathscr{AP})_N$. The following lemma is instrumental in establishing the relationship between problems $(\mathscr{P})$ and $(\mathscr{AP})_N$.

10.5. LEMMA. *The sequence $Q_N^{-1}u_N^*$ is bounded in U.*

Proof. The result is simple to establish if $\mathscr{U}$ is bounded (pick a sequence $\{v_N\}$ in $\mathscr{U}$ such that $Q_N v_N = u_N^*$; use equiboundedness of the linear functions $u \to Q_N^{-1}Q_N u$). Otherwise (cf. [2, Section 4]), choose an arbitrary $u \in \mathscr{U}$. Then for all N,

$$\Phi_N(u_N^*) \leqslant \Phi_N(Q_N u).$$

Lemmas 10.2 and 10.3 imply that the right-hand side converges to $\Phi(u)$. Thus, the sequence $\{\Phi_N(u_N^*)\}$ is bounded, which implies by (H4b)(ii) that $\{Q_N^{-1}u_N^*\}$ is also bounded. ∎

Let $\{u_{N(i)}^*\}_{i=1}^{\infty}$ denote a subsequence of $\{u_N^*\}$ having the property that the sequence $\{Q_N^{-1}u_{N(i)}^*\}$ is weakly convergent in U; let u^* denote its weak limit.

Since u^* is a natural candidate for a solution of problem $(\mathscr{P})$, we must consider whether it is an element of $\mathscr{U}$. Lemma 10.5 and hypothesis (H4) ensure that a bounded sequence $\{v_i\}$ exists in $\mathscr{U}$ with the property that $Q_{N(i)}v_i = u_{N(i)}^*$ for all i. Let v be the weak limit of the weakly convergent subsequence $\{v_{i(j)}\}$; since $\mathscr{U}$ is weakly closed, $v \in \mathscr{U}$. Lemma 10.4 implies that $Q_{N(i(j))}^{-1}u_{N(i(j))}^* \rightharpoonup v$. Hence $u^* \in \mathscr{U}$ because $Q_{N(i)}^{-1}u_{N(i)}^* \rightharpoonup u^*$ implies that $u^* = v$.

10.6. THEOREM. *The above u^* is a solution of problem $(\mathscr{P})$. Furthermore,*

$$\Phi(u^*) = \lim \Phi_{N(i)}(u_{N(i)}^*).$$

Proof. For all $u \in \mathscr{U}$ we have by Lemmas 10.1, 10.2 and 10.3

$$\begin{aligned}\Phi(u^*) &\leqslant \liminf \Phi_{N(i)}(u_{N(i)}^*)\\ &\leqslant \limsup \Phi_{N(i)}(u_{N(i)}^*)\\ &\leqslant \limsup \Phi_{N(i)}(Q_{N(i)}u)\\ &= \Phi(u). \quad ∎\end{aligned}$$

Under assumption of strict quasiconvexity, a standard argument involving subsequential limits implies that this u^* is the unique solution of problem $(\mathscr{P})$ and that $Q_N^{-1}u_N^* \rightharpoonup u^*$.

So far, we have not exploited the finite-dimensionality of $(\mathscr{A}\mathscr{P})_N$. Indeed, we shall not explicitly do so. This aspect of the approximate optimization problems is important, however, because it permits us to develop computational packages without introducing further approximations.

An example of a control set satisfying hypothesis (H4b)(iii) is given below. Control sets of this type, and those satisfying hypothesis (H4a), are commonly used in optimization problems.

Let X be a closed convex (unbounded) subset of R^m. Define the set

$$\mathscr{V} = \{u \colon u(t) \in X \text{ a.e. on } [a, b]\}.$$

Clearly, $\mathscr{V}$ satisfies hypothesis (H1). We employ a standard argument using the Hahn–Banach theorem to demonstrate that $\mathscr{V}$ satisfies hypothesis (H4b)(iii) as well. In particular, for $u \in \mathscr{V}$ and $i = 0, 1,..., \kappa N - 1$ let $y_i = y_i(u) = (N/r) \int_{K_i} u(t)\, dt$. Suppose that $y_i \notin X$ for some i. Then there exist $c \in R$, $\xi \in R^m$ such that $c < \xi^T y_i$ and $\sup\{\xi^T v \colon v \in X\} \leqslant c$. Therefore, $c < \xi^T y_i = (N/r) \int_{K_i} \xi^T u(t)\, dt \leqslant c$, a contradiction.

We now supply proofs for Lemmas 10.1, 10.2, 10.3 and 10.4.

Proof of Lemma 10.1. Our remarks in Section 8 imply that $z_N(u_N)$ factor converges to $z(u)$. Observe that for all $y \in E_N$,

$$\sup\{|(p_N^{-1}y)(t)|_Z \colon t \in [a, b]\} = |\, y\,|_{E_N}.$$

Using these facts and Lemma 4.2, we find that

$$\lim_N \sup\{|(p_N^{-1}z_N)(t) - z(t)|_Z \colon t \in [a, b]\} = 0.$$

Therefore $(p_N^{-1}z_N)(\zeta, \lambda, BQ_N^{-1}u_N)(b) \to z(b, a, \zeta, \lambda, Bu)$ and $x_N(\zeta, \lambda, BQ_N^{-1}u_N)(\cdot) \to x(\cdot, a, \zeta, \lambda, Bu)$ in Z and F respectively. The conclusion follows from the continuity of g_1 (hypothesis (H2)) and the weak lower semicontinuity of g_2 (hypotheses (H1), (H2) and Mazur's theorem). ∎

Proof of Lemma 10.2. This follows immediately from hypothesis (H2) and the convergence of $(p_N^{-1}z_N)(b)$ to $z(b)$ and $x_N(\cdot)$ to $x(\cdot)$ as above. ∎

Proof of Lemma 10.3. This may be proved in exactly the manner in which Lemma 4.1 was established. ∎

Proof of Lemma 10.4. For all $w \in U$ we have $\langle Q_N^{-1}Q_N v_N, w\rangle = \langle v_N, w\rangle - \langle v_N, w - Q_N^{-1}Q_N w\rangle$ since for $u_1, u_2 \in U$ we have $\langle Q_N^{-1}Q_N u_1, u_2\rangle = \langle u_1, Q_N^{-1}Q_N u_2\rangle$. The conclusion follows from Lemma 10.3. ∎

11. Description of Numerical Techniques Employed

In our discussion of numerical results, we shall not explicitly consider the effect of round-off error. Our intention is not to imply that it is negligible, but to emphasize the particular characteristics of the approximation scheme presented in Section 5. In this regard, note our comments in Section 13 concerning the chemostat example (for $N = 32$).

The finite difference scheme was used in the manner described below to compute solutions of problems $(\mathscr{A}\mathscr{P})_N$ (see Section 10). Its implementation was easy; in fact, the only further approximations required arose in the evaluation of $\pi_N\zeta$, $A(N, i, j)$ and $D(N, i, j)$ (see Sections 4, 5). A standard quadrature algorithm, exact for fourth-order polynomials, was employed for this purpose. Since $Q_N Q_N^{-1} w = w$ for each $w \in \mathscr{U}_N$, the set of admissible controls for each approximate problem was readily characterized.

The difference equations for the state and (see [13, Section VII.4]) auxiliary equations were solved exactly. Numerical solutions of problems $(\mathscr{A}\mathscr{P})_N$ were obtained by a combination of the gradient and conjugate gradient techniques. In particular, for a fixed N, a gradient step was taken on the first, fourth, seventh, etc. iteration, with conjugate gradient steps in between. This procedure was continued until a convergence criterion for the values of Φ_N was satisfied.

Several optimization problems governed by autonomous systems (namely examples 1, 2 and 3 of [3]) were solved numerically both by the technique described above, denoted as method R, and its analog (see [3]) for the method of Banks and Burns, denoted as method B. In each case the results were compared to an exact solution. As expected, low-order (i.e. $N = 4, 8$) approximations with method R were not very accurate. The results were better for intermediate values of N (i.e. 16, 32) and quite good for large N (i.e. 64, 128).

Generally speaking, comparable accuracy was obtained by taking N twice as large for method R as for method B. The definite, and not surprising, advantage of method R is its faster execution on a computer. Evaluating on this basis, methods B and R are roughly equivalent. Consider Example 4.3 of [3]: method B required about 32 seconds to execute with $N = 16$, while method R required 17 seconds with $N = 32$. Method B was only slightly more accurate.

12. Results for a Simple Nonautonomous System

We now present an example for which numerical and analytical solutions were readily obtained. These solutions were used to evaluate the accuracy of the above finite-difference scheme.

Consider the system

$$\dot{x}(t) = 6tx(t-1) + u(t) \qquad t \in [0, 2],$$
$$(x(0), x_0) = (1, 1),$$

and associated optimization problem ($\mathscr{P}$): minimize $\Phi(u) = (1/2)\, x^2(2) + (1/2) \int_0^2 u^2(t)\, dt$ over the set $\mathscr{U} = L_2(0, 2; R)$.

An analytical solution was obtained in the following manner by means of necessary conditions on extremal pairs (see [13, Section VII.2]). It is easy to verify that this problem is normal, so for an extremal pair (x^*, u^*) there is a function $\psi \in L_2(0, 3; R)$ such that

$$\psi(t) = \begin{cases} 0 & t \in (2, 3] \\ -x^*(2) & t = 2 \end{cases}$$

and on $[0, 2]$, $\dot{\psi}(t) = -6(t+1)\, \psi(t+1)$. Furthermore, u^* satisfies the pointwise maximum principle:

$$\begin{aligned} &\psi(t)\{6tx^*(t-1) + u^*(t)\} - (1/2)[u^*(t)]^2 \\ &\quad = \max_{v \in R} \{\psi(t)\{6tx^*(t-1) + v\} - (1/2)\, v^2\} \end{aligned}$$

Therefore $u^* = \psi$.

Letting $\psi(2) = \alpha$, we find that

$$u^*(t) = \psi(t) = \begin{cases} -3\alpha t^2 - 6\alpha t + 10\alpha & t \in [0, 1], \\ \alpha & t \in [1, 2]. \end{cases}$$

This in turn implies that

$$x^*(t) = \begin{cases} -\alpha t^3 + 3(1-\alpha)\, t^2 + 10\alpha t + 1 & t \in [0, 1], \\ (-1.2\alpha)\, t^5 + (4.5)\, t^4 + (26\alpha - 12)\, t^3 + (12 - 36\alpha)\, t^2 & \\ \qquad + \alpha t + (16.2\alpha - 0.5) & t \in [1, 2]. \end{cases}$$

Using the fact that $-\alpha = x^*(2)$, we obtain $\alpha = -23.5/44.8$.

Since we know a priori that problem ($\mathscr{P}$) has a unique solution (Section 9), these necessary conditions imply that it is given by (x^*, u^*) as defined above. Consequently the optimal cost Φ^* is given by

$$\begin{aligned} \Phi^* &= (1/2)[x^*(2)]^2 + (1/2) \int_0^2 [u^*(t)]^2\, dt \\ &= (1/2)\, 44.8\alpha^2 \\ &= (23.5)^2/89.6 \\ &= 6.1635. \end{aligned}$$

Selected values of Φ^*, u^* are given in Table I.

TABLE I

Nonautonomous Example

N	4	8	16	32	64	128	Exact
Φ_N^*	6.0509	6.0639	6.1050	6.1323	6.1474	6.1554	6.1635
CPU (sec.)	6	8	14	26	54	130	
Time	u_4^*	u_8^*	u_{16}^*	u_{32}^*	u_{64}^*	u_{128}^*	u^*
0.0	−5.0397	−5.1494	−5.1995	−5.2231	−5.2345	−5.2400	−5.2455
0.125		−4.6511	−4.7432	−4.7863	−4.8072	−4.8174	−4.8275
0.25	−3.7892	−4.1003	−4.2361	−4.2996	−4.3303	−4.3454	−4.3604
0.375		−3.4971	−3.6783	−3.7630	−3.8040	−3.8241	−3.8440
0.5	−2.3114	−2.8414	−3.0699	−3.1765	−3.2280	−3.2534	−3.2785
0.625		−2.1332	−2.4107	−2.5401	−2.6026	−2.6333	−2.6638
0.75	−0.6063	−1.3726	−1.7008	−1.8537	−1.9276	−1.9639	−1.9999
0.825		−0.5595	−0.9402	−1.1174	−1.2031	−1.2452	−1.2868
1.0	−0.6063	−0.5595	−0.5409	−0.5324	−0.5284	−0.5265	−0.5246

We have computed numerical solutions of the corresponding problems $(\mathscr{AP})_N$ for $N = 4$, 8, 16, 32, 64 and 128. A summary of the results is included in Table I. Observe that in this example the quantities $\pi_N(x(0), x_0)$, $A(N, i, j)$ and $D(N, i, j)$ were computed exactly. In particular,

$$\pi_N(x(0), x_0) = (1, 1, \dots, 1) \qquad \forall N,$$

$$A(N, i, j) = \begin{cases} 0 & j = 0 \quad \text{and all } N, i, \\ (6i + 3)/N & j = \nu = 1 \quad \text{and all } N, i, \end{cases}$$

$$D(N, i, j) = 0 \qquad \forall N, i, j.$$

The same general behavior is observed for this example as for the autonomous systems discussed above. In comparing the approximate solution for $N = 128$ to the analytical solution, we see that the relative error in the control values is less than 2%, except at time 0.75 where it is about 3%. The relative error in the payoff (for $N = 128$) is less than 0.2%.

13. The Chemostat—a Biochemical System

In this section we present numerical results for an optimization problem based on the dynamics of the chemostat. This device controls the growth of a microorganism population by regulating the supply of one essential nutrient, while providing all other nutrients in excess. Equations (13.1), (13.2) below are similar to models of the chemostat which have appeared in the literature (see [5], [17]). The nonlinearities in these equations are analogous to those appearing in the initial velocity approximations (usually associated with the names of Henri, Michaelis–Menton, and Briggs–Haldane) which are used to model enzyme catalyzed reactions (see [1, Ch. 1]).

Let x and s represent the microorganism population density and nutrient concentration in the growth chamber, respectively. Let the state of the system be given by $(x, s) \in C(-1, 3; R^2)$. Then we have

$$\dot{x}(t) = \frac{V_1 x(t) \int_{-1}^{0} s(t+\theta)\,\gamma(\theta)\,d\theta}{K + \int_{-1}^{0} s(t+\theta)\,\gamma(\theta)\,d\theta} - D(t)\,x(t) \qquad t \in [0, 3], \tag{13.1}$$

$$\dot{s}(t) = D(t)[s_0 - s(t)] - \frac{V_2 x(t)\,s(t)}{K + s(t)} \qquad t \in [0, 3], \tag{13.2}$$

where s_0 denotes the nutrient concentration in the incoming medium, $D(t)$ the washout rate, K the saturation constant for the rate of uptake of nutrient, V_1 the maximum growth rate, and V_2 the maximum uptake rate. The function $\gamma \in L_2(-1, 0; R)$ is used to weight the distributed effect of the nutrient concentration on the growth rate.

In the case to be considered below, we made no attempt to assign physically meaningful values to the above constants; rather, we arbitrarily set the values $s_0 = 2$, $K = 1$, $V_1 = 1$, $V_2 = 1/2$, and defined $\gamma(\theta) = 1$, $\theta \in [-1, 0]$.

The optimization problem $(\mathscr{P})$ was formulated as follows. A control $\tilde{D} \in L_2(0, 3; R)$ was employed to force the system away from an initial steady state. The system was then linearized about the resulting trajectory, denoted as $\tilde{Z} = ((\tilde{x}(t), \tilde{x}_t), (\tilde{s}(t), \tilde{s}_t)) \in C(0, 3; R^2 \times L_2(-1, 0; R^2))$. The cost function Φ: $\mathscr{U} = L_2(0, 3; R) \to R$ was defined as

$$\Phi(u) = 5 \int_0^2 |z(t)|^2\,dt + 50 \int_2^3 |z(t) - [\tilde{z}(2) - \tilde{z}(t)]|^2\,dt + 5 \int_0^3 |u(t)|^2\,dt + 50\,|z(3) - [\tilde{z}(2) - \tilde{z}(3)]|^2,$$

where $z = (x - \tilde{x}, s - \tilde{s})$ and $u = D - \tilde{D}$ represent the state and control functions respectively for the linearized system. This particular definition of Φ was motivated by the objectives:

(i) keep z and u small along the trajectory, so that the linearization is fairly accurate; and

(ii) force (x, s) to the new steady state $(x(t), s(t)) = (\tilde{x}(2), \tilde{s}(2))$, $t \in [2, 3]$.

We have computed numerical solutions for the corresponding problems $(\mathscr{AP})_N$ for $N = 4, 8, 16$. A summary of the results is included in Table II. The magnitude of $x_N^*(t), s_N^*(t)$ for $t \in [0, 2]$ and for each value of N is less than 0.1, except for $x_4^*(1.75)$ and $x_4^*(2.0)$. These magnitudes are less than 0.2.

TABLE II

Chemostat Model

N	4	8	16
Φ_N^*	0.2022	0.2283	0.2407
Time	u_4^*	u_8^*	u_{16}^*
0.0	0.0004	0.0673	0.0614
0.25	−0.0166	0.0077	0.0171
0.5	−0.0286	−0.0215	−0.0020
0.75	−0.0390	−0.0376	−0.0115
1.0	−0.0501	−0.0492	−0.0188
1.25	−0.0619	−0.0617	−0.0304
1.5	−0.0738	−0.0756	−0.0482
1.75	−0.0849	−0.0897	−0.0697
2.0	−0.1617	−0.1533	−0.1220
2.25	−0.1605	−0.1862	−0.1871
2.5	−0.1278	−0.1657	−0.1832
2.75	−0.0870	−0.1244	−0.1481
3.0	−0.0870	−0.1012	−0.1116
	$x_4^* - \tilde{x}$	$x_8^* - \tilde{x}$	$x_{16}^* - \tilde{x}$
2.0	0.1312	0.0991	0.0892
2.25	0.1022	0.0673	0.0565
2.5	0.0856	0.0490	0.0382
2.75	0.0761	0.0364	0.0237
3.0	0.0727	0.0268	0.0084
	$s_4^* - \tilde{s}$	$s_8^* - \tilde{s}$	$s_{16}^* - \tilde{s}$
2.0	−0.0686	−0.0434	−0.0225
2.25	−0.0538	−0.0278	−0.0056
2.5	−0.0455	−0.0200	0.0018
2.75	−0.0410	−0.0160	0.0047
3.0	−0.0400	−0.0147	0.0053

The runs displayed in Table II required much more time to execute than those appearing in Table I. A rough estimate (exact figures are not available) of the total CPU time required for $N = 4, 8, 16$ is 420 seconds. A run was made with $N = 32$. Numerical errors appear to have become significant here, for the CPU time required to meet the convergence criterion was greater than expected and the results differed slightly from what had been anticipated (based on an examination of earlier runs).

The rather lengthy execution times for this example show that batch processing is advisable in some cases. The program loaded into less than 256 K bytes of storage, and used five disk files with total length under 90 K bytes.

14. Concluding Remarks

The theory we have presented generalizes the work of Banks and Burns (see [2], [3]) on autonomous systems in two ways:

(i) nonautonomous systems are treated, and

(ii) the approximating systems are governed by difference, as opposed to differential, equations.

The primary practical advantage of (ii) is that it leads directly to algorithms which may be implemented on a computer, whereas the theory developed in [2], [3] requires a numerical approximation of the approximating ODE systems.

The use of averaging approximations in the study of hereditary systems is not at all new. For a detailed bibliography and commentary on the literature, we refer the reader to Section 5 of [3].

As mentioned earlier, the approximation scheme discussed in Section 5 has previously been studied. Delfour [7] investigated its convergence and applicability to the linear quadratic optimization problem. The techniques he employed are substantially different from those we have chosen. For purposes of comparison we shall use our notation to briefly describe his work. (The symbol Λ^0 below denotes a space similar to our Λ_∞.)

Delfour defines piecewise constant R^n-valued functions by means of the solutions of the difference equations obtained from the first component of Eqs. (5.2). These functions are shown [7, Proposition 3.2] to converge in the supremum norm to the solution of Eqs. (2.1), (2.2). Having asserted [7, Theorem 2.1] equivalence of Eqs. (2.1), (2.2) and the corresponding abstract ODE in Z for $(\zeta, \lambda, f) \in W \times \Lambda^0 \times F$, he restates [7, Proposition 3.4] the approximation results in operator notation in spaces similar to Z_N. Corresponding theorems are presented for solutions of the adjoint and Riccati equations, which lead to his treatment of the optimal control problem.

The major difference between Delfour's approach and that represented by [2], [3] and our own efforts is that (for the linear regulator problem) Delfour approximates not only the state equation in Z, but the infinite-dimensional adjoint and Riccati equations as well. The method of [2], [3] and this investigation involves immediate approximation of the state equation in Z by a finite-dimensional problem (either an ODE or difference equation) and then employment of standard numerical methods to solve the approximate problem.

REFERENCES

1. H. T. BANKS, "Modeling and Control in the Biomedical Sciences," Lecture Notes in Biomathematics, Springer-Verlag, Berlin/New York, 1975.
2. H. T. BANKS AND J. A. BURNS, An abstract framework for approximate solutions to optimal control problems governed by hereditary systems, *in* "Proceedings of the International Conference on Differential Equations, University of Southern California, September 1974" (H. A. Antosiewicz, Ed.), pp. 10–25, Academic Press, New York, 1975.
3. H. T. BANKS AND J. A. BURNS, Hereditary control problems: Numerical methods based on averaging approximations, *SIAM J. Control and Optim.*, in press.
4. H. T. BANKS AND A. MANITIUS, Projection series for retarded functional differential equations with applications to optimal control problems, *J. Differential Equations* **18** (1975), 296–332.
5. J. CAPERON, Time lag in population growth response of *Isochrysis galbana* to a variable nitrate environment, *Ecology* **50** (1969), 188–192.
6. J. W. DANIEL, "The Approximate Minimization of Functionals," Prentice-Hall, Englewood Cliffs, N. J., 1971.
7. M. C. DELFOUR, The linear-quadratic optimal control problem for hereditary differential systems: Theory and numerical solution, *Appl. Math. and Optim.* **3** (1977) 101–162.
8. N. DUNFORD AND J. T. SCHWARTZ, "Linear Operators," Part I, Interscience, New York, 1958.
9. L. M. GRAVES, "The Theory of Functions of Real Variables," McGraw-Hill, New York, 1956.
10. J. K. HALE, "Functional Differential Equations," Springer-Verlag, Berlin/New York, 1971.
11. S. G. KREIN, "Linear Differential Equations in Banach Space," American Mathematical Society, Providence, R. I., 1971.
12. S. LANG, "Analysis II," Addison-Wesley, Reading, Mass., 1969.
13. L. W. NEUSTADT, "Optimization: A Theory of Necessary Conditions," Princeton Univ. Press, Princeton, N. J., 1976.
14. A. PAZY, Semi-groups of linear operators and applications to partial differential equations, University of Maryland, Department of Mathematics, Lecture Note No. 10, College Park, Md, 1974.
15. D. C. REBER, "Approximation and Optimal Control of Linear Hereditary Systems," Ph. D. dissertation, Brown University, November 1977.
16. G. W. REDDIEN AND G. F. WEBB, Numerical approximation of nonlinear functional differential equations with L_2 initial functions, to appear.
17. T. F. THINGSTAD AND T. I. LANGLAND, Dynamics of chemostat culture: The effect of a delay in cell response, *J. Theor. Biol.* **48** (1974), 149–159.
18. G. F. WEBB, Functional differential equations and nonlinear semigroups in L^p-spaces, *J. Differential Equations* **20** (1976), 71–89.
19. K. YOSIDA, "Functional Analysis," 2nd printing (corrected), Springer-Verlag, Berlin/New York, 1966.